



XX Workshop de Investigadores en Ciencias de la Computación
26 y 27 de abril de 2018 - Corrientes - Argentina

LIBRO DE ACTAS

Facultad de Ciencias Exactas y Naturales y Agrimensura
Universidad Nacional del Nordeste

Red de Universidades con Carreras de Informática (RedUNCI)



Dapozo, Gladys

XX Workshop de Investigadores en Ciencias de la Computación - WICC 2018 : libro de actas / Gladys Dapozo ; Patricia Pesado ; compilado por Gladys Dapozo ; Emanuel Irrazabal. - 1a ed compendiada. - Corrientes : Universidad Nacional del Nordeste. Facultad de Ciencias Exactas, 2018.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-3619-27-4

1. Bases de Datos. 2. Minería de Datos. 3. Ingeniería de Software. I. Pesado, Patricia II. Dapozo, Gladys , comp. III. Irrazabal, Emanuel , comp. IV. Título.

CDD 004.071

Universidad Nacional del Nordeste (UNNE)

Rectora

Delfina Veiravé

Decana de FACENA

Lidia Itatí Ferraro

Autoridades Red de Universidades con Carreras de Informática (Red UNCI)

Coordinador Titular

Pesado Patricia (UNLP)

Coordinador Alterno

Estayno Marcelo (UNLZ)

Coordinadores de Área WICC 2018

Agentes y Sistemas Inteligentes

Marcelo Falappa (UNS)

Marcelo Errecalde (UNSL)

Daniel Pandolfi (UNPA)

Arquitectura, Redes y Sistemas Operativos

Luis Marrone (UNLP)

Daniel Arias Figueroa (UNSa)

Computación Gráfica, Imágenes y Visualización

Martín Larrea (UNS)

María J. Abásolo (UNLP-UNCPBA)

Roberto Guerrero (UNSL)

Ingeniería de Software

Pablo Fillotrani (UNSur)

Pablo Thomas (UNLP)

Fernanda Carmona (UNdeC)

Innovación en Sistemas de Software

Marcelo Estayno (UNLZ)

Guillermo Feierherd (UNTDF)

Osvaldo Sposito (UNLaM)

Gladys Dapozo (UNNE)

Procesamiento Distribuido y Paralelo

Marcelo Naiouf (UNLP)

Marcela Printista (UNSL)

Javier Balladini (UNCOMA)

Tecnología Informática aplicada en Educación

Zulma Cataldi (UBA-UTN)

Alejandra Zangara (UNLP)

Mónica Tugnarelli (UNER)

Gustavo Gil (UNSa)

Procesamiento de señales y Sistemas de Tiempo Real

Oscar Bría (INVAP)

Fernando Tinetti (UNLP)

Nelson Rodriguez (UNSJ)

Bases de Datos y Minería de Datos

Rodolfo Bertone (UNLP)

Claudia Deco (UNR)

Norma Herrera (UNSL)

Innovación en Educación Informática

Jorge Finocchietto (UCAECE)

Claudia Russo (UNNOBA)

Elena Durán (UNSE)

Seguridad Informática

Paula Venosa (UNLP)

Javier Echaiz (UNS)

Antonio Castro Lechtaller (IESE)

Comité Académico

UBA – Cs. Exactas	Garbervetsky, Diego	UBA – Ingeniería	Echeverría, Adriana
UN La Plata	Pesado, Patricia	UN Sur	Rueda, Sonia
UN San Luis	Piccoli, Fabiana	UNCPBA	Aciti, Claudio
UN Comahue	Vaucheret, Claudio	UN La Matanza	Spositto, Osvaldo
UN La Pampa	Alfonso, Hugo	UN Lomas de Zamora	Estayno, Marcelo
UNTierra del Fuego	Feierherd, Guillermo	UN Salta	Gil, Gustavo
UN Patagonia Austral	Lasso, Marta	UN SanJuan	Rodríguez, Nelson
UADER	Vivas, Patricia	UN Patagonia SJB	Buckle, Carlos
UN Entre Ríos	Tugnarelli, Mónica	UN Nordeste	Dapozo, Gladys
UN Rosario	Kantor Raul	UN Misiones	Kuna, Horacio
UNNOBA	Russo, Claudia	UN Chilecito	Carmona, Fernanda
UN Lanús	Rodríguez, Darío	UN Santiago del Estero	Duran, Elena
Esc. Sup. Ejército Arroyo	Arzubi, Alejandro	UN Litoral	Loyarte, Horacio
UN RioIV	Arroyo, Marcelo	UN Córdoba	Fridlender, Daniel
UN Jujuy	Herrera Cognetta, Analía	UN Rio Negro	Vivas, Luis
UN Villa María	Prato, Laura	UN Lujan	Panessi, Wálter
UN Catamarca	Poliche Maria Valeria	UN La Rioja	Campazzo Eduardo
UNTres de Febrero	Oliveros, Alejandro	UN Tucumán	Luccioni, Griselda María
UNAJ	Morales, Martín	UN Chaco	Austral Zachman Patricia
UN del Oeste	Foti, Antonio	UN de Cuyo	Forradelas, Raymundo
UN de Mar del Plata	Doumecq, Julio Cesar	U Morón	Padovani Hugo
UAI	De Vincenzi, Marcelo	U Belgrano	Guerci, Alberto
U Kennedy	Panizzi, Marisa	U Adventista del Plata	Bournissen Juan
UCAECE	Finocchietto, Jorge	U Palermo	Alvarez Adriana
UCA Rosario	Grieco, Sebastián	U Salvador	Zanitti, Marcelo
U Aconcagua	Giménez, Rosa	U GastónDachary	Beyersdorf, Carlos
UCEMA	Guglianone, Ariadna	U Austral	Cassol, Ignacio
U Atlántida Argentina	Rathmann, Liliana	UCA La Plata	Bertone, Rodolfo
ITBA	Mon, Alicia	U Champagnat	Pinciroli, Fernando
UN Hurlingham	Medrano Gustavo		

Comité Organizador FACENA UNNE

Gladys Noemi Dapozo

Cristina Greiner

David La Red Martinez

Emanuel Irrazabal

María Viviana Godoy Guglielmone

Sonia Itatí Mariño

Coordinadores Locales Por Área

Agentes y Sistemas Inteligentes

Sonia Itatí Mariño.

Arquitectura, Redes y Sistemas Operativos

Leopoldo José Rios.

Base de Datos y Minería de Datos

David La Red Martinez.

Computación Grafica, Imágenes y Visualizaciones

Raquel Petris.

Ingeniería de Software

María Ferraro y Oscar Vallejos.

Innovacion en Educación en Informática

Cristina Greiner.

Innovacion de Sistemas de Software

Emanuel Irrazábal.

Procesamiento Distribuido y Paralelo

Alejandro Burgos.

Procesamiento de Señales y Sistemas de Tiempo Real

Jorge Veglia.

Seguridad Informatica

Ricardo Monzón.

Tecnologia Informática Aplicada en Educación

Beatriz Castro Chans.

Indice de Artículos

Agentes y Sistemas Inteligentes	1
Desarrollo de Habilidades sociales y cognitivas a través de Sistemas Inteligentes	2
Operadores de Cambio de Creencias No-Priorizada en Sistemas Argumentativos Aplicado para Programación Lógica Rebatible	7
Aplicación de técnicas de Inteligencia Artificial para evaluar la producción científico-académica de investigadores de Universidades públicas del Noroeste Argentino	12
Diseño de Algoritmos Inteligentes Aplicados a Interfaces Humano-Computador e Internet de las Cosas	17
Monitoreo de carga por métodos no invasivos en el hogar argentino utilizando redes neuronales	22
Deep Neural Network para Análisis Acústico	27
Modelo basado en la Toma Decisiones con Criterios Múltiples para la elección de metodologías de Data Science	32
Diseño de Método de Ensamble Homogéneo para Clasificadores Débiles usando un esquema de reducción de datos simultaneo basado un enfoque co-evolutivo.	37
Aprendizaje Automático, Aplicaciones en Visión por Computadora	42
Redes Inteligentes de Agua: Factores y métodos para la predicción del consumo residencial de agua potable	46
Soluciones inteligentes para el desarrollo urbano sostenible	50
Metaheurísticas aplicadas de problemas de Scheduling con restricciones	56
Análisis automático de grandes volúmenes de datos en redes sociales mediante minería de textos combinado con algoritmos inteligentes	60
Integración de bases de creencias manteniendo coherencia y consistencia con criterio unificado de credibilidad	65
Optimización a gran escala usando metaheurísticas	70
Desarrollo de Sistemas de Scheduling para la Nube	75
Modelo de Inteligencia Artificial aplicados al Trading Algorítmico	80
GA2LA: Grupo de Aprendizaje Automático, Lenguajes y Autómatas	84
Sistemas de análisis textual en formato no estructurado	89
Procesamiento de textos estructurados	94
Arquitectura, Redes y Sistemas Operativos	98
Análisis de tráfico Multicast de video H264/Theora en Redes Wi-Fi IEEE 802.11ac	99
Sistemas de comunicación robustos para infraestructuras avanzadas de medición de energía implementadas sobre PLC (Power Line Communications)	104
Nuevo Modelo de Decisión para Gestión de Tráfico en Redes	109
Metodología para la selección de recursos computacionales gestionados con tecnologías de cloud computing en ambientes educativos	114
Nueva propuesta para la administración de recursos y procesos en sistemas distribuidos	118
Visualización Gráfica de resultados de Simulaciones de Redes de Sensores Inalámbricos	123

Implementación de sistemas de control automático para cultivos hidropónicos en invernaderos de la provincia de Misiones	128
El problema de las comunicaciones rurales: estudio y selección de las mejores soluciones	133
Análisis de Protocolos de Comunicaciones para Internet de las Cosas	138
Hub Of Things: Concentrador para el Internet de las Cosas	142
Análisis de Eficiencia y Tolerancia a Fallos en Arquitecturas Multiprocesador para Aplicaciones de Procesamiento de Datos	147
Tecnologías de la información y las comunicaciones mediante IoT para la solución de problemas en el medio socio productivo	153
Contienda entre las variantes del protocolo TCP Vegas y Reno por los recursos de la red en un modelo híbrido simple	159
Evaluación de performance en Redes Definidas por Software	164
Servidor virtual para detección de intrusos y ataques en IPV6	169
Puerta de Enlace para Internet de las Cosas usando Computadora Industrial Abierta	174
Predicción y alerta temprana de incendios forestales mediante integración de WSN e imágenes satelitales en un método de reducción de incertidumbre guiado por datos	179
CIAA en la industria misionera: Fabrica de ladrillos. Integración con software SCADA y ERP	184
Rendimiento de Aplicaciones Web en plataformas de contenedores de código abierto	189
Bases de Datos y Minería de Datos	194
Modelo de Decisión para la Validación de Métodos de Imputación Mediante la Utilización de Algoritmos de Minería de Datos	195
Reconocimiento de Patrones Genéticos por medio de Grafos	200
Calidad Universitaria mediante Técnicas del Data Mining	205
Avances en selección de biomateriales utilizados en implantes dentales aplicando técnicas de minería de datos	209
Estructuras de Datos y Algoritmos Eficientes para Búsquedas Web y Procesamiento de Grandes Datos	214
Procesamiento Eficiente de Grafos Masivos para Aplicaciones en Redes Sociales	220
NoSQL: Modelos de datos y sistemas de gestión de bases de datos	225
Infraestructura de Acceso a Datos Primarios con Aporte de Semántica en Repositorios Digitales	229
Sistemas Recomendadores aplicados en Educación	234
Benchmarking de Bases de Datos NoSQL para el almacenamiento de Modelos Semánticos	238
Visualización en Ciencia de Datos	242
Análisis de similitud en documentos de texto mediante técnicas de ciencia de datos basadas en aprendizaje profundo (Deep Learning)	246
Construcción de una arquitectura BI para optimización de las decisiones a partir del estudio del nivel de madurez de los entornos tecnológicos empresariales	251
Recuperación de la Información	255
Estudio y análisis de técnicas de modelado de grandes volúmenes de datos jurídicos	261
Diseño y Construcción de Procesos de Explotación de Información para el Área de Ciencias de la Computación	265
Hacia la optimización del uso de datos abiertos en el ámbito público	270
Técnicas de unificación de datos para la visualización de grandes volúmenes de datos	275
Bases de Datos y Aplicación de Técnicas de Avanzadas	280
Aplicación de minera de datos para facilitar el tratamiento de las normas de producción de atributos semánticos en idioma español	285
Tecnologías de la Web Semántica aplicadas al tratamiento de documentos jurídicos electrónicos	290

Extracción de conocimiento en redes sociales mediante herramientas de software libre y plataformas de hardware paralelo-distribuidas	295
Implementación de técnicas específicas de minería de datos en aplicaciones web con motores de base de datos relacionales	300
Visualización en un entorno de minería de texto	306
Propuesta de procedimiento para el análisis delictivo basado en la explotación de la información . . .	311
Indexación y Búsqueda sobre Datos no Estructurados	316
Algoritmos de aprendizaje automático para respuestas en tiempo real sobre entornos masivos de datos	321
Minería de Datos y Visualización de Información	325
Polimorfismos de Nucleótidos Simples Relacionados al Riesgo de Enfermedades: Clasificación Automática de Estudios Epidemiológicos de Tipo Caso-Control Utilizando Técnicas de Minería de Texto	330
Métodos de Acceso para Bases de Datos Métricas	335
Buenas prácticas en el diseño de estructuras de datos en bases de datos relacionales	340
Recuperación de Datos e Información en Bases de Datos Masivas	345
Minería de Datos y Big Data. Aplicaciones en riesgo crediticio, salud y análisis de mercado	350
Computación Gráfica, Imágenes y Visualización	355
Líneas de Investigación del Grupo de Realidad Aumentada Aplicada: Templates de Catálogos Aumentados Integración Escalable de Realidad Aumentada basada en Imágenes y Rostros Aumentación de Sistemas SCADA en el Contexto de la Industria 4.0	356
Métricas, Técnicas y Semántica para la Visualización de Datos	361
Tecnologías Inmersivas Aplicadas: Realidad Virtual y Aumentada	366
Plataforma Detección de objetos en tiempo real	371
Realidad Aumentada y Visión por Computador. Framework multipropósito	376
Medición automática de pulso cardíaco utilizando imágenes de video	384
Aplicaciones de Visión por Computador, Realidad Aumentada y TVDi	389
Realidad Virtual y Realidad Aumentada como medios para un lenguaje generativo multimodal	396
Reconocimiento de patrones de imágenes digitales obtenidas mediante microscopio y parametrizadas según la técnica de Micronúcleos y la técnica Ensayo Cometa empleada por el Laboratorio de Citogenética General y Monitoreo Ambiental UNaM-IBS-CONICET para la detección de daños celulares.	401
Reconstrucción y animación 3D	406
Innovación en Educación en Informática	411
Los Entornos Virtuales Flexibles para el Desarrollo de Competencias en el Área de Algoritmos y Lenguajes de Programación	412
Sistema Experto con actualización de reglas en cascada para la construcción de indicadores en Prácticas de Educación Digital y Robótica Educativa.	417
Estrategias innovadoras en los Procesos de Enseñanza y de Aprendizajes de Informática	420
Influencia del uso de la Gamificación y las herramientas de evaluación continua en el proceso de enseñanza y aprendizaje	425
Integración de servicios en aplicación móvil para aprendizaje ubicuo	430
Sistema de producción Polimedia FA.C.E.N.	435
Herramientas de interpretación gráfica para codificación de diagramas de modelado de sistemas interpretables por un disminuido visual	441
Procesos de diseño de componentes para espacios virtuales de trabajo orientados a la educación de personas con discapacidades	446

Pensamiento Computacional y Programación en la Formación de Docentes del Nivel Primario	451
Entorno basado en programación por bloques para robots educativos	456
Experiencias educativas, reflexiones, aportaciones docentes e investigaciones de la praxis en el contexto universitario	461
Entorno Web para la Programación del Robot Frankestito	465
El lugar de las Ciencias de la Computación en el currículum de la Escuela Secundaria Argentina . . .	470
La Universidad y los procesos de transformación: el Género en las TIC. El caso de la Facultad de Informática de la UNLP	475
Analizando el impacto de talleres de programación en escuelas con respecto al ingreso de alumnos en carreras de informática	480
Estrategias innovadoras de enseñanza de la programación y didácticas específicas para fomentar el pensamiento computacional	484
EscuelasTIC: las tecnologías digitales en las aulas	489
La tecnología como vehículo de articulación Nivel Medio / Universidad	494
Edimbrujo: Hacia la definición de un modelo didáctico para la enseñanza de la Inteligencia Artificial en Juegos	500
Metodología para la definición y ponderación de factores de éxito para procesos de gestión de proyectos académicos unipersonales de práctica profesional supervisada en carreras de informática . .	505
Ingeniería del Software	510
Requisitos de usuario y gestión de la demanda en AOP4ST	511
Un Modelo de Análisis para aplicación de Patrones de buenas prácticas en el Modelado Conceptual de Gobierno Electrónico	516
Evaluación Temprana de la Usabilidad Empleando Patrones Embebidos en la Construcción del Modelo Conceptual para Aplicaciones Web	521
Análisis de Framework web adaptativos basados en HTML5	526
Bases de Datos de Proyectos de Software Ágiles para ser Utilizadas en Simulación de Proyectos. . . .	531
Movilidad y Gestión del Tráfico: Automatización y Aplicación de un Modelo de Calidad para Flotas Dinámicas en una Ciudad Inteligente	537
Hacia un Modelo de Calidad de Software como soporte a los procesos de licitación en el ámbito estatal	542
Smart City - Modelo de calidad mixto para software de gestión de movilidad urbana	547
Ingeniería de software para sistemas embebidos, requisitos en PYMEs y Testing continuo	553
Proponiendo un enfoque integrador para diseñar y evaluar interfaces de usuario web	558
Ingeniería de software para sistemas críticos ferroviarios	563
Desarrollo e implementación de un modelo basado en estándares para el logro y evaluación de la Usabilidad en aplicaciones informáticas	568
Aplicación Reflexiva de un Proceso de Requisitos	573
Evolución de los Factores Situacionales durante el Proceso de Requisitos	578
Evaluación de Compatibilidad y Complejidad para el Reuso de Servicios	583
Tendencias en el desarrollo de Aplicaciones para Dispositivos Móviles	588
Gobernanza digital. Mejora de procesos de gestión y calidad de software	593
Metodologías, técnicas y herramientas de ingeniería de software en escenarios híbridos	599
Coordinación de dispositivos ubicuos: una solución basada en SOA y coreografías	604
Enfoques de optimización multi-objetivo basados en preferencias en la ingeniería de software	609
Integración de IoT en el modelado de procesos de negocio	614
Revisión con base cognitiva de un Proceso de Requisitos	619
Estrategias para la clasificación de contenido y usuarios de Foros de Discusión Técnicos	624

Generación automática de API REST a partir de API Java, basada en transformación de Modelos (MDD)	629
Arquitectura de Software en el Proceso de Desarrollo Ágil. Una Perspectiva Basada en Requisitos Significantes para la Arquitectura.	635
Planificación de las Pruebas del Software	640
Evaluación de Variantes de Inspección en la Ingeniería de Requisitos	645
Modelos de Madurez para la Mejora de Calidad de los Indicadores de Desarrollo Sostenible	650
Los Procesos de Negocio en las Smart City: Un nuevo Paradigma	655
Un Framework para la Evaluación de Modelos Conceptuales de Procesos de Negocio	660
Implementación de una arquitectura de procesos como resultado de la aplicación del ciclo de vida BPM durante sus fases de configuración y ejecución	665
Proceso de auditoría para micro, pequeñas y medianas organizaciones	670
Aproximaciones para el desarrollo multiplataforma de Aplicaciones Móviles	675
Aplicación móvil multiplataforma sensible al contexto para el turismo religioso en Santiago del Estero	680
Un Método de Evaluación para Modelo de Gestión de Calidad en Empresas de SSI	685
Inclusión de Hacking ético en el proceso de Testing de software	690
Accesibilidad Web, aportando a la inclusión	694
Diseño de software basado en BPM para mejorar la usabilidad de las aplicaciones	698
Gestión del conocimiento y sistemas informáticos. Una propuesta para las organizaciones del siglo XXI	703
Fases para el desarrollo de programas paralelos y su aplicación en la predicción del comportamiento de incendios forestales	706
Modelo de Sistema de Gestión de Calidad para un Instituto de Educación a Distancia	710
Modelo para la interoperabilidad entre controladores de Redes Definidas por Software.	716
Metodologías y herramientas emergentes para contribuir con la calidad del software	721
Modelo de interoperabilidad entre sistemas para trazabilidad de procesos internos de la industria cárnica empleando nuevas tecnologías de identificación	726
Línea de Investigación en Evaluación de Productos Software	731
Diseño y construcción de sistemas de IOT seguros y escalables	737
Análisis y Gestión de Riesgo en Proyectos Software. Un nuevo modelo integrando la metodología SEI y Magerit2	742
Innovación en Sistemas Software	747
Análisis de nuevos estándares para la web móvil	748
Implementación de soluciones con Raspberry Pi accesibles desde dispositivos móviles	752
Interfaces tangibles - Una nueva forma de interactuar con los Smartphones	756
Innovación en la gestión de la salud asistencial y laboratorios de investigación en salud	760
Servicios PaaS de cloud computing e informática contextualizada aplicada a una app híbrida, turística para la ciudad de Puerto Iguazú Misiones	765
La información científica en la UNdeC: ¿Aumento la visibilidad web?	770
Uso de VRPN en la implementación de una BCI para rehabilitación neurológica	776
Una Herramienta Gráfica con Razonamiento basado en DL para el Análisis de Modelos OVM	780
Influencias del estado biométrico-emocional de personas interactuando en contextos de entornos virtuales	785
Aplicaciones móviles 3D	790
Evaluación del desarrollo tecnológico para la definición de Industrias 4.0	794
Hacia la Formalización de un Lenguaje Visual Unificador de UML, EER y ORM2	799
Juego de Realidad Aumentada para Incentivar la Actividad Física en Niños con Discapacidad	804
Algoritmos de Layout Automático para una Herramienta Multi-Vistas de Modelado Ontológico	809

Hacia un Lenguaje Gráfico para SPARQL-DL	814
Auditoría automatizada de Trazabilidad Vitivinícola	819
Tratamiento de evidencias digitales forenses en dispositivos móviles	824
Juegos serios y aplicaciones interactivas usando realidad aumentada y realidad virtual	829
Robótica basada en internet de las cosas	835
Propuesta de recolección de datos para aforo vehicular	838
Utilización de tecnología RFID/NFC para el desarrollo de un sistema de llenado selectivo de envases. Industria 4.0.	843
Realidad Virtual y Aumentada, Big Data y Dispositivos Móviles: Aplicaciones en Turismo	848
Dispositivo de asistencia de personas mediante monitoreo IoT	854
Gestión del control de acceso con tecnología Open Source en proyectos de domótica	859
Tecnología aplicada al deporte de alto rendimiento	864
Data-driven Agriculture, las tecnologías aplicadas a problemas de agricultura regional	869
Técnicas y estrategias para evaluar las ideas en la plataforma IdTR.	875
Plataforma de Google: Android Mobile como soporte de Android Things	881
Rampas Digitales Innovativas para Personas con Discapacidad	886
Agentes Inteligentes y Web Semántica: Verbalización en una herramienta Web de modelado ontológico	891
Procesamiento Distribuido y Paralelo	896
Orquestación de servicios para la Continuidad Edge al Cloud	897
Modelado y Simulación de Sistemas de Gran Escala	902
Procesamiento Computacional Paralelo con Metaheurísticas Híbridas para la Reducción de Incertidumbre en Modelos de Incendios Forestales	907
Performance de arquitecturas multiprocesador: técnicas de modelado y simulación en HPC y Cloud Computing	912
Performance de cloud computing para HPC: despliegue y seguridad	917
Sintonización automática de aplicaciones paralelo/distribuidas basadas en algoritmos evolutivos y evolución diferencial	922
Garantizando la consistencia de JavaScript en un contexto de memoria compartida	927
Simulación computacional, ciencia de los datos, cómputo de alto rendimiento y optimización aplicados a mejorar la predicción de modelos de simulación que representan la evolución de sistemas complejos.	930
Arquitecturas Multiprocesador en HPC: Software, Métricas, Modelos y Aplicaciones	935
Cómputo paralelo y distribuido para HPC. Fundamentos, Construcción y Evaluación de Aplicaciones.	941
Tendencias en Arquitecturas y Algoritmos para Sistemas Paralelos y Distribuidos	947
Aplicaciones de Cómputo Intensivo con Impacto Social	953
Estrategias y análisis orientados al manejo de datos masivos usando computación de alto desempeño	958
Procesamiento de Señales y Sistemas de Tiempo Real	963
Modelos y Métodos Computacionales Aplicados I	964
Energía Eficiente	968
Robótica y redes de sensores en sistemas de tiempo real	974
Estación Terrena de Adquisición de Señales de Satélites no Geoestacionarios	978
Sistema de comunicación para la implementación de redes inalámbricas de sensores	982
Sistemas Embebidos y de Tiempo Real con Planificación Heterogénea	987
Visión artificial aplicada en Agricultura de Precisión	992
Sensores: Pilares de la investigación tecnológica	997

Seguridad Infomática	1002
Voto Electrónico Seguro con Criptografía Homomórfica1003
Criptografía Liviana para aplicar en dispositivos IoT1008
Entorno de disponibilidad forense para la recolección de datos digitales en HTTP1011
Evaluación de la Seguridad en Sistemas Informáticos1016
Protección de activos vinculados con la información - preparación para la Ciberdefensa1021
Confidencialidad de las comunicaciones en Sistemas Móviles1026
Toolkit de Criptoanálisis1029
Análisis de Vulnerabilidades de Sistemas Web en desarrollo y en producción1033
Vulnerabilidades en HTTP/21038
Análisis del Marco Normativo Legal para el Ciclo de Vida de la Evidencia Digital1043
Software Abierto para la Evaluación de Sistemas Criptológicos Integrados1047
Técnicas y Herramientas para Regular la Seguridad en Web Services Basados en WSDL1051
Investigación en ciberseguridad: un enfoque integrado para la formación de recursos de alto grado de especialización1056
Seguridad en la Nube: Almacenamiento de Imágenes Médicas y Watermarking1061
Análisis Sistemático de la Seguridad en Internet of Things1066
Tecnología Informática Aplicada en Educación	1072
Dispositivos móviles como instrumento facilitador del aprendizaje1073
Determinación de perfiles de rendimiento académico en la UNNE con Minería de Datos Educativa1078
Ambientes de Aprendizaje enriquecidos con Tecnología1083
Modelos y herramientas para el proceso de desarrollo de Serious Games1088
Plataforma Gamificada que permite la resolución de desafíos en lenguaje C en el ámbito universitario1093
Diseño, modelado, simulación e implementación de técnicas emergentes de teoría de juegos en Serious Games1098
Analíticas de Aprendizaje para Serious Games1102
Vinculación con docentes de nivel secundario y con ingresantes de Informática, a través de talleres y juegos educativos desarrollados, uno de los cuales, es un juego móvil.1106
Realidad virtual en terapias para estimular la cognición de niños con encefalopatía crónica no evolutiva1111
Aplicación de las Reglas de Usabilidad en Sitios Web Universitarios1115
Modelo predictivo para la determinación de riesgo de deserción universitaria para ingresantes de la UTN FRLP1120
UAICASE: Enseñanza de Ingeniería de Software en Entornos Virtuales Colaborativos1124
Procedimientos y herramientas para la mejora de indicadores académicos en facultades de ingeniería1130
Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos1135
Tendencias en interacción personas-ordenador. Aplicaciones en escenarios educativos1141
Posibilidades de los dispositivos móviles para la educación en colegios de nivel secundario1146
Vinculación de líneas de Investigación y desarrollo con la extensión universitaria1151
Marco de trabajo para la evaluación de usabilidad de objetos de aprendizaje basados en Realidad Aumentada1156
Computación Aplicada en Ciencias y Educación1161
Realidades Alternativas como Herramientas de Mediación para el desarrollo del Pensamiento Computacional1166
Repositorios Digitales con Contenidos Orientados a las Necesidades de Instituciones Educativas de la Región1171

Análisis y detección de patrones en un grafo conceptual construido a partir de respuestas escritas en forma textual a preguntas sobre un tema específico1176
Caracterización preliminar de los aspirantes a ingresar a la Licenciatura en Ciencias de la Computación	181
Implementación de un EVEA institucional para enriquecer la enseñanza de pregrado, grado y posgrado de la UNTDF1186
Diseño y Aplicación de Estrategias para la Enseñanza Inicial de la Programación1192
REFORTICCA: Recursos para el Empoderamiento de FORMadores en TIC, Ciencias y Ambiente1197
M-learning con Realidad Aumentada basada en Objetos 3D1203
Tecnologías aplicadas a educación en UNNOBA1207
Realidad aumentada y otras tecnologías emergentes en procesos de enseñanza y aprendizaje.	
Aproximaciones metodológicas al diseño y evaluación de propuestas didácticas1212
Recursos TIC para el aprendizaje de la química y la física en el ciclo básico universitario1217
Aplicación de técnicas estadísticas y de minería de datos para el análisis de perfiles de rendimiento académico: El caso de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de la Pampa1222
Desarrollo de una Plataforma de Capacitación a Distancia Utilizable y Abierta para Personas con Discapacidad Visual1227
Modelo de sistema tutor inteligente de apoyo en la realización de proyectos de investigación en posgrados en ingeniería1232
Uzi: Máquina virtual sobre plataforma Arduino para robótica educativa - Resultados de la primer etapa	236
Análisis de aspectos de diseño en cursos masivos en línea. Caso de estudio sobre un curso de accesibilidad web1240
Análisis y desarrollo de una herramienta para el manejo de los operadores del Álgebra Relacional . .	.1246
Evaluación de calidad de herramientas utilizadas en el enseñanza de la programación basada en ISO 250001249
Tesis Doctorales	1254
Agent SocialMetric: herramienta de asistencia al docente en el dominio Educativo del aula1255
Estudio de la influencia de un entorno de simulación en la enseñanza de redes de computadoras en el nivel universitario1265
Optimización de rendimiento, justicia y consumo energético en sistemas Multicore asimétrico mediante planificación1275
Un enfoque inteligente para la selección de grupos de expertos mediante redes sociales1286
Reconocimiento de gestos dinámicos y su aplicación al lenguaje de señas1296
Formalización y Generalización del Manejo de Preferencias en Servicios de Razonamiento Rebatible	.1306

Autores

A

ABÁSULO María José (UNLP): 356, 1141, 1197, 389
ACOSTA Julio (UNNE): 1078,118, 195
ACOSTA Karina B. (UNaM): 330
ACOSTA M. R. (UNCa): 1246
ACOSTA Nelson (UNCPBA): 417, 804, 401
ADO Mariana (UNNOBA): 1207
AGOSTINI Federico (UNNE): 118
AGUILAR Santiago (UNTDF): 1192
AGUILERA María Belén (UNTDF): 1186
AGUILERA Sergio Omar (UB): 731
AGUIRRE Sofía (UNLPAM): 1073
AGUIRRE Verónica (UNLP): 599
AHMAD Tamara (UNCPBA): 1207
ALBORNOZ María Claudia (UNSL): 480
ALDERETE Claudia (UNLaM): 526
ALDERETE Romina Y. (UNNE): 694
ALFONSO Hugo (UNLPAM): 70
ALFONZO Pedro L. (UNNE): 694
ALLAN Claudia (UNCOMA): 1161
ALONSO Daniel (UNAJ): 153
ALONSO DE AR Ana (UNCOMA): 1161
ALONSO José Manuel (UNNE): 553
ALONSO Nancy (UNSJ): 75
ALONSO Nicolás (UNNOBA)(UNCPBA): 997, 1207
ALTOAGUIRRE Paula (UNLP): 886
ALVARADO Y. (UNSL): 396
ALVAREZ Eduardo (UNNOBA): 997, 992
AMADEO Ana Paola (UNLP): 1240,475
AMATRIAIN Hernán (UNLA): 441,690,80
AMBROGIO Ezequiel (UTN-FRC): 1061
AMBROSI Viviana (UNLP): 614
AMIGONE Federico (UNCOMA): 500
ANABALON Diego (UNCOMA): 583
ANFUSO Daniela Lobos (UNCa): 1246
ANGELERI Paula (UB): 731
ANTONIO Ramiro (UNLP): 42
APARICIO Florentina (UNJu): 17
APARICIO María (UNJu): 568
APOLLINI Ruben (UNSL): 902
ARAGÓN Fabiana (UNJu): 568
ARAMAYO Fernando Rúben (UNJu): 12, 300
ARANDA Gabriela (UNCOMA): 624
ARANDA ROMERA Juan Antonio (UNSJ): 635

ARCH Daniel (UTN-FRC): 1061
ARCHUBY Federico (UNLP): 1141
ARES Alicia E. (UNaM): 209
ARÉVALO Sergio (UNdeC): 494
ARGUELLO Santiago (UTN-FRC): 340
ARIAS Silvia (UTN-FRC): 1061
ARIEL Giannone (UNLA): 690
ARMANDO Silvana (UTN-FRSFco): 542,655
ARROYUELO Jorge (UNSL): 335
ARRÚA Martín (UTN-FRCU): 609
ARSAUTE Ariel (UNRC): 629
ARTOLA Verónica (UNLP): 1141, 1135
ARZUBI Alejandro Arroyo (UNDEF): 133
ASCHERI María Eva (UNLPAM): 1146
ASTUDILLO Gustavo Javier (UNLPAM): 420
AUBIN Verónica I. (UNLaM): 425
ÁVILA Federico Carrilao (UNLP): 1240
AZAR Miguel Augusto (UNJu): 12,17
AZAR Paola (UNSL): 316
AZCURRA D. (UNLA): 835

B

BACINO Gustavo Alfredo (UNMdP): 1083
BADARACCO Numa (UNNE): 721
BAEZA Natalia (UNCOMA): 1161
BAIGORRIA Lorena S. (UNSL): 660
BALDASSARRI Sandra (UZ): 1141
BALLADINI Javier (UNCOMA): 953, 947
BALLESTEROS Carlos (UNLPAM): 1073,1227
BALMER Agustín (UNNOBA): 992
BANCHERO Santiago (UNLu): 220,321
BANCHIO Leandro (UTN-FRSFco): 726
BANCHOFF TZANCOFF Claudia M. (UNLP): 829, 489
BAQUINZAY Manuel (UNCa): 1246
BARBERO Juan (UNLPAM): 70
BAREIRO Hernán (UGD): 123, 531, 184
BARONE Mirta (UNLaM): 255
BARRIONUEVO Diego (UTN-FRC): 1033
BARRIONUEVO Mercedes (UNSL): 958
BARRIOS Walter G. (UNNE): 461
BASGALL María José (UNLP): 941,350
BAST Silvia (UNLPAM): 420
BATTAGLIA Nicolás (UAI): 1124
BATTISTA Anabella De (UTN): 316
BAZÁN Patricia (UNLP): 614
BECERRA Martín Ezequiel (UNLaM): 356

BEDETTI Nicolás (UNLaM): 578
 BEGUERÍ Graciela (UNSJ): 242, 306
 BEINOTTI Raúl (UTN-FRSFco): 22
 BELCASTRO Ángela (UNPSJB): 1106
 BELIZÁN Maximiliano (UNAJ): 912, 917
 BELLONI Edgardo (UGD): 531, 184
 BELOSO Juan Pablo (UNNOBA): 997
 BELTRAMONE Matías (UTN-FRSFco): 22
 BENDATTI Natalia (UNNOBA): 1207
 BENITEZ Juan De Dios (UGD): 123, 531, 330, 184
 BERMÚDEZ Carlos (UNLPAM): 70
 BERNAL Rubén Alfredo (UNNE): 563
 BERNARDI Emanuel (UTN-FRSFco): 978
 BERNARDIS Edgardo (UNSL): 1051
 BERNARDIS Hernán (UNSL): 1051
 BERÓN Mario Marcelo (UNSL): 1051, 881, 1066, 737
 BERTONE Rodolfo (UNLP): 1073,1106,599
 BEZZONE Eber (UNLPAM): 70
 BIANCHINI Germán (UTN-FRM): 907, 922, 179
 BIANCIOTTI Andres (UTN-FRSFco): 726
 BIANCO Santiago (UNLA): 441, 446, 80
 BIASIN Estefano (UAI): 968
 BINDA Adriana (UNSE): 824
 BJERG Ernesto (UNS): 366
 BLANCO Claudio (UNTDF): 1186
 BLANCO Sebastián (UAI): 927,1236
 BLAUTZIK Leonardo (UNLaM): 425
 BOLATTI Diego (UTN-FRRe): 164
 BOND Román (UNAJ): 912
 BORACCHIA Marcos (UNLP): 593
 BOSSERO Julio C. (UNLaM): 255
 BOUCIGUEZ María José (UNCPBA): 1197
 BOYLE Bruno (UAI): 968
 BOZA Ramiro (UNLP): 389
 BRACCO Luciano (UTN-FRCU): 609, 325
 BRACHETTA Mariana (UTN-FRM): 430
 BRAUN Germán (UNCOMA): 780, 799, 809, 814, 891
 BREA Ricardo (UK): 142
 BRIA Oscar (UNLP): 964
 BRITOS Luis (UNSL): 345
 BRITOS Paola (UNRN): 32
 BRUNO Javier (UTN-FRSFco): 22
 BRUNO Pablo (UNCOMA): 953
 BUCKLE Carlos (UNPSJB): 229
 BUDÁN Paola (UNSE): 675
 BUENO Matías (UTN-FRC): 340

BUSANO Facundo (UTN-FRSFco): 978
 BUSTOS AGUIAR M. Soledad (UNCa): 1246
 BUSTOS Miguel Alfredo (UNSL): 881, 1066

C

CABALLERO Javier (UTN-FRM): 819
 CABALLERO Sergio (UNaM): 742
 CABRAL Cristian (UGD): 184
 CABRERA José Luis (UNLaM): 425
 CACERES Nélide Raquel (UNJu): 238
 CAFFETTI Jacqueline Diana (CONICET): 401
 CAFFETTI Yanina Andrea (UNaM): 401
 CALABRESE Julieta (UNLP): 593
 CALCAGNO Ricardo (UTN-FRRe): 164
 CALIBAR Andrea Belén (UNSJ): 246
 CALLONI Juan (UTN-FRSFco): 726, 716
 CAMILETTI Pablo (UNLPAM): 1146
 CAMMARO Pablo (UNLaM): 752
 CANDIA Damián (UNLP): 614
 CÁNDIDO Andrea (UNJu): 568
 CANTERO Alejandro Lucas (UNaM): 265
 CANTEROS Ana Valeria (UNaM): 265
 CAPECCI Magali (UNLP): 599
 CAPPELLETTI Marcelo Angel (UNAJ): 153
 CARBALLO L (UNPA): 50
 CARBALLO Nicolás (UM): 179
 CÁRDENAS Adrián (UTN-FRM): 99
 CARDENAS Marina E. (UTN-FRC): 89,94, 84
 CARDOZO Claudia (UNPA): 558
 CARLOS Francisco Jorge (UNJu): 17
 CARMONA Fernanda Beatriz (UNdeC): 770,133,1171,494,869
 CARNUCCIO Esteban (UNLaM): 854
 CAROD Nadina Martínez (UNCOMA): 624
 CARRANZA Alvaro (UNSE): 1203, 675
 CARRASCO Agustín (UTN-FRC): 340
 CARRIZO A. (UNPA): 558
 CARRIZO Blanca Rosa (UTN-FRC): 1115
 CARRIZO Claudio (UTN-FRSFco): 542, 270, 655
 CARRIZO Marcelo (UM): 179
 CASALI Ana (UNR): 451
 CASAMAYOR Agustín (CAECE): 2
 CASANOVA Belén (UNCOMA): 953
 CASANOVA Carlos (UNSL): 280
 CASANOVA Carlos (UTN-FRCU): 609, 325
 CASATTI Martín (UTN-FRC): 37, 1176
 CASERES German (UNLP): 588

CASTAÑEDA Sebastián (UNLP): 389
 CASTILLO Julio (UTN-FRC): 84, 89, 94
 CASTRO Alicia (UNSL): 902
 CASTRO Claudia (UTN-FRC): 521
 CASTRO Franco (UNSJ): 242
 CASTRO Leandro (UNLPAM): 420
 CASTRO LECHTALER Antonio (UNdeC): 770
 CASTRO Marcelo (UNJu): 568
 CASTRO María L (UNCPBA): 1197
 CASTRO Silvia (UNS): 361,366, 406
 CATANIA Carlos (UNCUYO): 1181
 CATENA Franco (UTN-FRM): 670
 CAVALLERA Giselle (UNSL): 537
 CAVERO Lorena Verónica (UNLPAM): 1222
 CAYLÁ Ivana Belén (UM): 179
 CAYMES-SCUTARI Paola (UTN-FRM): 907, 179
 CAZÓN Liliana (UNJu): 568
 CEBALLOS Jorge (UB): 731
 CECCHI Laura (UNCOMA): 465, 780, 799, 809, 814, 456
 CECHICH Alejandra (UNCOMA): 583, (UNCOMA): 624
 CELESTE Carolina (UNCOMA): 1161
 CENICH Gabriela (UNCPBA): 1197
 CERRUTI Santiago (UNC): 27
 CERVINO Claudio (UM): 785
 CESPEDES Eugenia (UNLaM): 640,384
 CHARCZUK Norberto (UNLA): 446
 CHARNE Javier (UNNOBA): 997
 CHARNELLI M. Emilia (UNLP): 234
 CHÁVEZ Edgar (CICESE): 335, 345,
 CHAVEZ Susana (UNSJ): 897
 CHEEIN David (UCSE): 680
 CHICHIZOLA Franco (UNLP): 935, 941, 947
 CICCO Carlos Di (UNNOBA): 997
 CICERCHIA Benjamín (UNNOBA): 992
 CIPRIANO Marcelo (EST-IESE)(UNDEF)(USAL): 138, 1026, 1029, 1047, 1008
 CLUSELLA María M (UCSE): 680
 COBIALCA Silvia (CAECE): 285
 COCCONI Diego (UTN-FRSFco): 22, 665
 COIMBRA Rodrigo Pereyra (UNaM): 553
 COLACIOPPO Nicolás (UTN): 1115
 COMPANY Ana (UNNE): 698, 484, 721
 CONDO Sergio (UNRN): 1212
 CONSTABLE Leticia (UTN-FRC): 84
 CONTRERAS Micaela (UTN-FRM): 838
 CORBALÁN Leonardo (UNLP): 588
 CORDOBA Diego (UM): 179
 CORGATELLI Franco (UNCOMA): 583
 CORNEJO María Eda (UNCOMA): 470
 CORONEL Gabriel (UCP): 189
 CORSO Cynthia (UTN): 1176, 37
 CORTÉS Lucía (UNCUYO): 1181
 CORTÉS María Alicia (UNNE): 760
 CORTEZ Alberto (UTN-FRM): 819
 CORTEZ Marcos Manuel (UNCOMA): 470
 CORVI Julieta (UNLP): 350
 COSTA Verónica Gil (UNSL): 902
 COTO Facundo (UTN-FRCU): 325
 CRISTALDO Patricia (UTN-FRCU): 325
 CRISTINA Federico (UNLP): 790
 CRUZ Alejandro (UNdeC): 804, (UNdeC): 494
 CUEVAS Carlos (UTN-FRRe): 164
 CUEVAS Juan Carlos (UTN-FRC): 1033
 CUEVAS Verónica (UNCOMA): 1212
 CUITIÑO Alfonso (UNLP): 588
 CUNEO Juan Manuel (UNSJ): 75, 875

D

DAMIANO Luis Esteban (UTN-FRC): 340
 DAMONTE Tomas (UB): 645
 DANIELE Marcela (UNRC): 629
 DAPOTO Sebastián (UNLP): 790
 DAPOZO Gladys (UNNE): 1156, 698, 484, 721, 1249
 DAPOZO ROMERO Ramiro (UNLP): 864
 DASSO Aristides (UNSL): 1016
 DE BATTISTA Anabella (UTN-FRCU): 609, (UTN-FRCU): 325
 DE BRITTO E SILVA Erik (UFMG): 716
 DE GIUSTI Armando (UNLP): 389, 947, 953,935, 941, 964, 974
 DE GIUSTI Laura (UFRGS): 935, 941, 947
 DE LA FUENTE Juan (UNCOMA): 456
 DE LUCA Graciela (UNLaM): 854
 DE MARIA Eduardo (UNLaM): 794
 DE RENZIS Alan (UNCOMA): 583
 DE VITTO Cecilia (UNNOBA): 1207
 DEAGUSTINI Cristhian (UNS): 65
 DECO Claudia (UNR): 345
 DEGIOVANNI Federico (UTN-FRSFco): 716
 DEL GIORGIO Horacio (UNLaM): 794, 425
 DELGADO Andrea (UTN-FRC): 521
 DELIA Lisandro (UNLP): 588, 848

DELL OSO Matías (UNLP): 935
 DELVECHIO Tomás (UNLu): 214
 DEPETRIS Beatriz (UNTDF): 848, 1186, 1192
 DESTEFANIS Eduardo A. (UTN-FRC): 27
 DI GENARO María E (UNSL): 335
 DI GIONANTONIO Alejandra (UTN-FRC): 1115,
 1033, 1061
 DIAZ Alfredo Ariel (UNJu): 12
 DÍAZ Daniel (UNSJ): 75, 875
 DÍAZ Javier (UNLP): 234, 1011, 614, 829, 1038, 1240,
 475, 864, 1056, 886
 DÍAZ Mario (UNSJ): 412
 DIBEZ Pablo (UNPSJB): 1106
 DIESER María Paula (UNLPAM): 1222
 DIESTE Oscar (UPM): 604
 DIEZ E. (UNLA): 80
 DISTEL Juan M. (UNLPAM): 420
 DOMINGUEZ Facundo (UNaM)(UNLaM): 209, 843
 DOMÍNGUEZ Juan (UNTDF): 1192
 DOORN Jorge (UNLaM): 573, 578, 619
 DORZÁN María Gisela (UNSL): 280
 DOVIS Emmanuel (UTN-FRSFco): 978
 DUARTE David (UNAJ): 917

E

ESCALANTE Jaquelina E. (UNNE): 1232
 ECHAZÚ Alejandro Luis (IESE): 133
 ECKERT Karina (OTRA): 290
 ECKERT Karina (UNaM): 32
 ELGUETA Rodrigo Atilio (UM): 179
 ELIZALDE María Celia (UNO): 578, 619
 ENCINAS Diego (UNAJ)(UNLP): 912, 917, 974, 147,
 935
 ERRECALDE Jorge (UNLP): 50
 ESCARZA Sebastián (UNS): 361
 ESCUDERO Damián (UNSL): 335
 ESNAOLA Leonardo (UNCPBA): 60
 ESPÍNDOLA María Cecilia (UNNE): 484, 1249
 ESPINOSA Agustín (UTN-FRM): 819
 ESPONDA Silvia (UNLP): 593
 ESTEVEZ Carlos (OTRA): 1047
 ESTEVEZ Elsa (UNS): 650
 ESTRADA Bruno Rafael (UNJu): 17
 ESTREBOU César (UNLP): 42
 ESZTER Edgardo (UM): 785
 ETEROVIC Jorge E. (UNLaM)(USAL): 1043, 1008,
 138

EVANS Felipe (UNMdP): 1088

F

FACCHINI Higinio (UTN-FRM): 99
 FALAPPA Marcelo A. (UNS): 65
 FARFÁN José (UNJu): 300, 568
 FAROPPA Martín (UNNOBA): 992
 FAVA Laura (UNLP): 829, 864, 886
 FAVRET Fabián (UGD): 330, 184
 FEIERHERD Guillermo (UNTDF): 848, 1186, 1192
 FELISSIA Sergio Francisco (UTN-FRSFco): 978
 FENNEMA Marta C. (UNSE): 675
 FERNÁNDEZ Alberto (CETINIA): 547
 FERNANDEZ BARIVIERA Aurelio (Universitat
 Rovira i Virgili): 350
 FERNÁNDEZ Darío (UNDEF): 1047
 FERNANDEZ David (UNNOBA): 1207
 FERNANDEZ Ezequiel E. (UTN-FRC): 521
 FERNANDEZ G. (UNLA): 835
 FERNANDEZ GIL Alberto (URJC): 537
 FERNANDEZ J. (UNSL): 396
 FERNANDEZ Jacqueline (UNSL): 1166
 FERNÁNDEZ Juan Manuel (UNLu): 321
 FERNANDEZ Mirta (UNNE): 461
 FERNANDEZ S. (UNLA): 835
 FERNÁNDEZ TAURANT Juan Pablo (UTN-FRC):
 521
 FERRARI Mariano A. (UNPSJB): 987
 FERRARINI Cintia (UNSJ): 1111
 FERRARO M. (UNNE): 721
 FERREYRA Juan Pablo (UTN-FRSFco): 665
 FEUERSTEIN Esteban (UBA): 220, 214
 FIEGE Lucas (UGD): 123
 FIERRO Ariel (UNTDF): 1192
 FIGUEROA Daniel Arias (UNSa): 1265, 710
 FIGUEROA Karina (UMSNH): 335
 FIGUEROA Liliana María (UNSE): 824
 FIGUEROLA Claudio (UNLaM): 794
 FILIPPI Jose Luis (UNLPAM): 1073, 1227
 FILLOTTRANI Pablo Rubén (UNS): 780, 799, 809,
 814, 650
 FLORES Andrés Pablo (UNCOMA): 583
 FLORES Lorena (UNNE): 311
 FLORES Luis Rojas (UNPSJB): 350
 FLORES Sebastián (UNSL): 737
 FONSECA E SILVA Vinicius (UFMG): 716
 FONSECA Efraín (EST-IESE): 604

FORNAROLI Mauro F. (UNER): 1011, 1038
FORTIN Daniel (UM): 785
FOTI Antonio (UNTREF): 133
FRACCHIA Carina (UNCOMA): 1161
FRATI Fernando Emmanuel (UNLP)(UNdeC): 941,
770, 1171, 494
FRATI José Nicolás (UNdeC): 494
FRUTOS Mariana (UNRC): 629
FUNES Ana (UNSL): 1016
FUNKNER Sofía (UNLPAM): 1222
FUSARIO Rubén Jorge (IUE): 133

G

GAETAN Gabriela (UNPA): 558
GAGLIARDI Marisa (UNdeC): 1171
GAGLIARDI Olinda Edilma (UNSL): 280
GALAIN Cintia (UNNE): 694
GALARZA Brian (UNAJ): 917
GALLO Silvana Lis (UNLP): 941
GANUZA María Luján (UNS): 361
GANZ Nancy (UNaM): 209, 169
GARABETTI Miguel Mendez (CONICET)(UM):
922, 706, 179
GARCÍA Edith (UNDEF): 1026,1029,1047
GARCÍA Federico (UNLaM): 843
GARCÍA Mariano (UTN-FRM): 819
GARCÍA Mario Alejandro (UTN): 27
GARCIA Gerardo (UNLaM): 854
GARCÍA Pablo (UNLPAM): 1003
GARCÍA Roberto (UNLA): 446
GARINO Carlos Garcia (UNCUYO): 1181
GARIS Ana (UNSL): 480
GARRIDO Matías (UNCOMA): 891
GARRIGA Martín (UNCOMA): 583
GASTAÑAGA Iris (UTN-FRC): 1033
GAUDIANI Adriana Angélica (UNGS): 930, 941
GAZCÓN Nicolás Fernando (UNS): 366
GETTE Marina (UNSL): 685
GIACOMANTONE Javier (UNLP): 964
GIBELLI Tatiana (UNRN): 1212
GIBELLINI Fabian (UTN-FRC): 1033
GIL Gustavo D. (UNSa): 710
GIMENEZ Christian (UNCOMA): 814
GIMENEZ Juan (UNPSJB): 1106
GIMENEZ Liliana (UNNE): 1217
GIMENEZ Mónica (UNLAR): 897
GIOIA Cintia (UNLaM): 1043

GIORDANO Andrés (UNLu): 220
GIORDANO Luis A (UNLu): 321
GIORLANDO Agustín (UM): 179
GIULIANELLI Daniel (UNLaM): 854
GODOY Diego Alberto (UGD): 123, 531, 330, 184
GODOY G. María Viviana (UNNE): 1232, 703, 461
GOITEA Oscar Alejandro (UNLaM): 425
GÓMEZ Sebastián (OTRA): 953
GOMEZ Soledad (UNLP): 489
GONZALEZ Alejandro (UNLP): 1135
GONZALEZ Ariel (UNRC): 629
GONZÁLEZ Federico (UNTDF): 848
GONZÁLEZ Mónica (UNSJ): 1111
GONZALEZ Valeria A. (UNSa): 710
GONZÁLEZ Verónica (UNaM): 128
GONZÁLEZ Claudia N. (UNPA): 159
GORGA Gladys (UNLP): 1135, 1141
GOURIC Guillermo Adrián (UNSJ): 295
GRAMAJO Sergio (UTN): 164
GREINER Cristina (UNNE): 484, 721, 1249
GRIN Germán Kurt (UBA): 1171
GROPPO Mario A. (UTN-FRC): 516
GROSSO Alejandro (UNSL): 335
GUAREPI Valentín (UNAJ): 153
GUASCH María Mercedes (UNNOBA): 1207
GUATELLI Renata (UNLaM): 573, 425
GUERRERO R. (UNSL): 1166, 396
GUEVARA Andrea (UTN-FRC): 340
GUEVARA Miguel (UNSJ): 897
GUTIÉRREZ Laura V. (UNSJ): 306
GUZMÁN Analía (UTN): 1176

H

HADAD Graciela (UB)(UNLaM)(UNO): 645, 578,
619
HADERNE Marisa Fabiana (UNCUYO): 1181
HARARI Ivana (UNLP): 1240, 886
HASPERUÉ Waldo (UNLP): 350
HAURECH Hugo (UNaM): 114
HENRIQUES Pedro R. (UMinho): 881, 1066, 737
HERNANDEZ Nicolás (UTN-FRC): 1061, 89
HERRERA Myriam (UNSJ): 205
HERRERA Norma Edith (UNSL): 316, 325
HERRERA Susana (UCSE): 680, 1203, 675
HIDALGO Fabian (UTN-FRM): 99
HINOJAL Hernán (UNMdP): 1083
HOLLEGER Jorge (UNSJ): 246

HORAS Fabiola (UNTDF): 848
HUBBE Susana (UTN-FRM): 819
HUERTAS Francisco (UNTDF): 848

I

IBAÑEZ Eduardo (UNLP): 599, 1135
IBAÑEZ Francisco (UNSJ): 75, 875
IERACHE Jorge (UM): 785
IERACHE Jorge (UNLaM): 356, 1156
IGARZA Santiago (UNLaM): 356, 1043
IGLESIAS Luciano (UNLP): 1135, 935
IGLESIAS Pedro (UNNOBA): 1207
ILLESCAS Gustavo (UNCPBA): 1130
INTINI Aníbal (UNDEF): 133
INZA Leonardo (UB): 376
IRRAZÁBAL Emanuel (UNNE): 760, 553, 563, 698, 721
IRURZUN Ivana (UNSE): 1203
ISAIA Claudia Pura Elena (UNdeC): 494
ISTVAN Romina Mariel (UTN FRLP): 1120

J

JARA Jimena (UNAJ): 912
JASZCZYSZYN Adrian (UNNOBA): 997
JATIP Nicolás (UNNOBA): 997
JIMBO SANTANA Patricia (UCE): 350
JOFRÉ N. (UNSL): 396
JOSELEVICH María (UNAJ): 153

K

KAPLAN Gladys (UNLaM): 573, 640
KASIÁN Fernando (UNSL): 345
KIMURA Isabel (UNLP): 489
KLENZI Raúl Oscar (UNSJ): 246, 295, 306
KOGAN Pablo (UNCOMA): 500
KORNUTA Cristian (UNaM): 765
KRUIOSKI S. (UNaM): 265
KUCUK Lucas G. (UNaM): 1156
KUHNS Franco (UNMdP): 1102, 1088
KUMVICH Augusto (UNLaM): 843
KUNA Horacio (UNaM): 209, 128, 265, 742, 505
KUNYSZ Eduardo (UNAJ): 147
KURT GRIN Germán (UBA): 133
KUZ Antonieta (UNCPBA): 1255

L

LA FRAZIA L. (UNLP): 42
LA RED MARTINEZ David Luis (UNNE): 104, 118, 1078, 195, 109, 114
LAFUENTE Guillermo Javier (UNLPAM): 1073, 1227
LAFUENTE Gustavo (UNLPAM): 1073
LANFRANCO Einar Felipe (UNLP): 1056
LANZARINI Laura (UNLP): 234, 42, 350
LARA Cecilia (UNSE): 824
LARREA Martín (UNS): 361, 366, 406
LARREGUI Juan I. (UNS): 406
LASAGNA Valeria (UTN FRLP): 1120
LAURENT Roberto (UNCOMA): 1161
LAVALLÉN Pablo (UNLu): 214
LECHTALER Antonio Castro (UBA)(UNDEF): 1171, 133, 1021, 1026, 1029
LEDESMA Viviana (UNLaM): 578
LEGUIZAMÓN Guillermo (UNSL): 46, 56
LENCINA Paula (UNNOBA): 1207
LEÓN Oscar (UTN-FRM): 430
LESCA Norma (UCSE): 261, 275, 824
LESTANI Juan (UB): 376
LETURIA Fernando (UCSE): 680
LEZCANO Juan M. (UGD): 290
LIBUTTI Leandro (UNLP): 935, 947
LIGORRIA Karina (UTN): 1176
LIGORRIA Laura (UTN-FRC): 1115
LIPORACE Julio (UNDEF): 1026, 1029, 1047
LLAMPA Álvaro (UNJu): 300
LLANOS Emanuel (UNNOBA): 997
LLARENA Myriam (UNSJ): 412
LOBOS Martín Mariano (UNLPAM): 1146
LOMBARDO Graciela (UNaM): 128
LOOR Fernando (UNSL): 902
LOPEZ Gabriel Vicente (IESE): 1047
LÓPEZ Luis (UNLaM): 200
LOPEZ Mariano Emanuel Alejandro (UNNE): 859
LÓPEZ Pedro (UAI): 968
LOPRESTI Mariela (UNSL): 958
LORENTI Luciano (UNLP): 964
LORENZO Martín (UB): 376
LOVOS Edith (UNRN): 1212
LUCERO Irene (UNNE): 1217, 484
LUCERO Maximiliano (UNSL): 958
LUCERO Walter A (UNSL): 685
LUDUEÑA Verónica (UNSL): 335, 345

LUENGO Pablo (UNNOBA): 992, 997
LUJÁN Arnaldo (UM): 179
LUNA Carla (UNLP): 42
LUND María Inés (UNSJ): 205, 225
LUPI Daniel (UNLaM): 843
LUQUE Claudio (UTN-FRC): 37
LUQUE Emilio (UAB): 930
LUQUE FADÓN Emilio (UAB): 953

M

MACIA Nicolás (UNLP): 1056
MADOZ Maria Cristina (UNLP): 1135
MAIDANA Josue (UNNE): 694
MAIORANO Ariel (UNDEF): 1026, 1029, 1047
MALBERNAT Lucia Rosario (CAECE): 251, 285
MALBERTI RIVEROS María Alejandra (UNSJ):
242, 306
MALDONADO Calixto (UTN-FRC): 37, 340
MALDONADO Marilena (UNSE): 1203
MALLEA Daniel Aguil (UNTDF): 1192
MALVACIO Eduardo (EST-IESE)(UNDEF): 1047,
1026, 1029
MAMANI Arnaldo Ismael (UNJu): 17
MAMANI Jonatan (UNTDF): 1192
MAMANI Nahuel (UNCOMA): 1161
MANERA Roxana (UTN-FRC): 1115
MANGIARUA Nahuel Adiel (UNLaM): 356
MANINO Gustavo (UTN-FRM): 819
MANRESA YEE Cristina (UIB): 1141
MANRIQUES Patricia (UNdeC): 494
MANSILLA Gladis Alejandra (UNLPAM): 1073
MARCHISIO Susana (UNNE): 1232
MARCISZACK Marcelo (UTN-FRC): 516, 521
MARCOVECCHIO Ignacio (UNS): 650
MARÍA Victoria Santa (UBA): 200
MARIANETTI Osvaldo (UM): 776
MARINELLI Giuliano (UNCOMA): 809
MARINELLI Marcelo (UNaM): 128, 169
MARIÑO Sonia (UNNE): 311, 694, 461, 703, 1232
MARKO Isabel B. (UNLaM): 526
MARRERO Luciano (UNLP): 599
MARRONE Agustín (UNLu): 214, 321
MARRONE Luis (UNLP): 159
MARTÍN Adriana (UNPA): 558
MARTÍN Adriana (UNSJ): 897
MARTÍN Ariel (UTN-FRM): 819
MARTIN María Cristina (UNLPAM): 1222

MARTIN Sofía (UNLP): 829, 475
MARTINEZ Carlos (UTN-FRM): 819
MARTINEZ Cintia Valeria (UNSL): 335
MARTÍNEZ Gimena (UTN-FRC): 37
MARTINEZ M. Vanina (UNER): 65
MARTINEZ María ROxana (CAETI): 1124
MARTINEZ Mariano (UNCPBA): 1130
MARTÍNEZ Roxana (UAI): 748, 756
MARTINI E. (UNaM): 265
MARTINS Adair (UNCOMA): 1161
MARTINS Sebastian (UNLA): 311, 441, 690, 80
MASCHERONI Agustín (UNNE): 721, 553
MASSA Stella Maris (UNMdP)(CAECE): 1083, 1088,
1102, 2
MAURICIO Nordio (UNLP): 1141
MAZZANTI Renato (CENPAT): 229
MAZZEO Hector (UTN FRLP): 982
MEDEL Diego (UNSJ): 897
MEDINA Oscar (UTN-FRC): 516
MEDINA Santiago (UNLP): 974
MEDINA Yanina (UNNE): 721
MEDRANO José Federico (UNJu): 12, 17
MENDEZ GARABETTI Miguel (CONICET)(UCH):
907, 897
MENDEZ Mariano (UNLP): 974, 935
MENNA Alvaro Castro (UM): 785
MENVIELLE María Alejandra Paz (UTN): 1176
MERCADO Gustavo (UTN-FRM): 174
MERCADO Jorge N (UNSJ): 897
MERCADO Viviana (UNPA): 50
MERCEDES COR Irma Rafaela (UNJu): 300
MERENDA Franco (UNSL): 345
MERLINO Hernan (UNLA): 446, 690, 80
MESÍA Natalí Salazar (UNLP): 1141
MICHELAN Gastón (UNCOMA): 500
MIGANI Silvina (UNSJ): 225
MIGHETTI Juan (UNLaM): 578
MINETTI Gabriela (UNLPAM): 70
MIRANDA Andrea (UNCPBA): 1197
MIRANDA Natalia (UNSL): 958
MIRETTI Marco (UTN-FRSFco): 978
MITARITONNA Alejandro (CITEFA)(UB): 389, 376
MOGRO Nelson Ariel (UNJu): 300
MOJELAT Natalia (UNR): 451
MOLINA Ana Laura (UNSJ): 706
MOLINA D. (UNPA): 50
MOLINA Roberto (UNLP): 350

MOLINARI Lia (UNLP): 475, 1056
 MOLLOJA Josue Emanuel (UNJu): 17
 MON Alicia (UNLaM): 794
 MONETTI Julio (UTN-FRM): 838, 430
 MONTE Cristian Perez (UNSL): 958
 MONTEJANO Germán (UNSL): 1003, 1016, 604, 660, 716
 MONTENEGRO Cristian (UNPA): 50
 MONTES DE OCa Erica (UNLP): 941, 947
 MONTES Leonardo (UNJu): 300
 MONTEZANTI Diego (UNAJ)(UNLP): 147, 935
 MONZON Francisco Tonin (UNLu): 214, 321
 MORALEJO Lucrecia (UNLP): 1141
 MORALEJO Raul (UTN-FRM): 371, 670
 MORALES María I. (UNSE): 1203
 MORALES Martin (UNAJ): 147, 912, 917, 153
 MORÁN Marina (UNCOMA): 953
 MORAN Matias (UNNOBA): 997
 MORAN Ricardo (UAI): 927, 1236
 MORCELA Antonio (UNMdP): 1088
 MORELLI José (UM): 179
 MORENO Edgardo J. (UNLaM): 1093
 MORENO Juan Carlos (UTN-FRC): 521
 MORENO Marcelo P. (UNSJ): 635
 MORESI Alejandro (UNDEF): 1021
 MORETTI Nicolas (UNNOBA): 1207, 997
 MORO Lucrecia Ethel (UNMdP): 1083
 MOUNIER Monica (UNaM): 330
 MOYANO Marcelo (UNCOMA): 583
 MOYANO Miguel (UTN-FRC): 27
 MUGAS Martín Navarro (UTN-FRC): 1061
 MULASSANO Micaela (UTN-FRSFco): 22
 MULKI Juan (UCSE): 261
 MUÑOZ Mercedes (UTN-FRM): 430
 MUÑOZ Roberto (UTN-FRC): 1033, 340
 MUÑOZ Rocío (UNLP): 593
 MURAZZO María (UNSJ): 897, 706, 958
 MUSCIA Alex (UNNOBA): 997

N

NAIOUF Marcelo (UNLP): 930, 935, 941, 947, 389
 NAJAR RUIZ Pablo J. (UNSE): 675
 NATALA Matías (UNNOBA): 997
 NAVARRO Martín (UTN-FRC)(UTN-FRM): 89, 838
 NAVARRO Mirta Elizabeth (UNSJ): 635
 NAVEDA Claudia (UTN-FRM): 819
 NEIL Carlos (UAI): 1124

NICOLET Santiago (USAL): 138
 NUÑEZ Juan Pablo (UTN-FRCU): 325

O

OCHIPINTI Pedro (UNNOBA): 1207
 OLARIAGA Sandra (UTN-FRC): 1115
 OLGUIN Luis Alberto (UNSJ): 1111
 OLGUÍN María G. (UNPSJB): 987
 OLIVA Elisa S (UNSJ): 875
 OLIVAS VAREL Jose Angel (UCLM): 350
 OLIVEROS Alejandro (UNTREF): 133
 OLMEDO Silvina (UB): 376
 OLSOWY Verena (UNLP): 599
 OMAR BIALE Claudio (UNaM): 265
 ONEDDU Eduardo (UNaM): 169
 OPORTO POBLE Pablo Gabriel (UNdeC): 494
 ORLANDI Cristina (OTRA): 953
 ORMEÑO Emilio (UNSJ): 1111
 OROZCO Javier D. (UNPSJB): 987
 OROZCO Sergio Daniel (UNPA): 50, 56
 ORTEGA Manuel (UNSJ): 295, 306
 ORTEMIN SÁNC Facundo Nehuén (UNCOMA): 624
 ORTIZ Claudia S. (UNLu): 640, 384
 ORTIZ DE ZARATE Juan M. (UBA): 220
 OSELLA MASSA Germán Leandro (UNNOBA): 997
 OSIO Jorge R. (UNAJ): 147, 153, 917
 OSORIO Maria Alejandra (UNLP): 1240, 475
 OTAZÚ Alejandra (UNSJ): 875
 OTOYA BET Raúl (UM): 776
 OVIEDO Sandra (UNSJ): 75, 875
 OYARZUN Ángela (UNCOMA): 780

P

PABLO RODRÍG Ismael (UNLP): 935
 PACÍFICO Cristian Damián (UNER): 65
 PACINI Elina (UNCUYO): 1181
 PAEZ Carolina (UB): 376
 PÁEZ Francisco E. (UNPSJB): 987
 PAEZ Sergio (UTN-FRSFco): 716, 726
 PAGANO Matias (UNLP): 864
 PALAVECINO Rosa (UNSE): 1203, 675
 PALLIOTTO Diana (UNCa): 1246
 PALMERO Pablo Rafael (UNSL): 280
 PALOMO Lilia (UCSE): 261, 275
 PANDOLFI Daniel (UNPA): 46, 56, 50
 PANESSI Walter F. (UNLu): 640, 384

PANIEGO Juan Manuel (UNLP): 974, 935, 947
 PANIZZI Marisa (UK): 142
 PAPINI María Cecilia (UNCPBA): 1197
 PAREDES Julio César (UNJu): 300
 PAREDES Rodrigo (UTALCA): 335
 PARISI Germán (UTN-FRC): 1033
 PARRA BELTAN Fernando A. (UAI): 748
 PARRA Lorena Vanesa (UNSJ): 635
 PARRA Susana (UNCOMA): 1161
 PASCAL Andrés (UTN): 316
 PASCUAL Kevin Isaías (UNCOMA): 1161
 PASINI Ariel (UNLP): 593, 599
 PAUTSCH G. (UNaM): 265
 PAZ MENVIELLE María Alejandra (UTN-FRC): 521
 PECCHIA Matías (UTN-FRM): 174
 PEÑASCO Andres (UM): 179
 PENDENTI Horacio (UNTDF): 1186, 1192
 PEPE María Laura (UNLaM): 573
 PERALTA Lorena (UTN-FRC): 1115
 PERALTA Mario (UNSL): 537, 547, 270, 655, 660, 685, 542
 PEREIRA Maria Joao (IPB): 1051
 PERESSINI Emiliano (UNdeC): 770, 1171
 PERETTI Gaston (UTN-FRSFco): 978
 PERETTI Juan Pablo (UTN-FRC): 340
 PEREZ Daniel (UNLPAM): 1073
 PEREZ Marisa (UTN-FRSFco): 665
 PEREZ Matías (UNdeC): 670, 494, 869
 PÉREZ Nicolás (UNLA): 441, 80
 PEREZ Nicolás (UTN-FRC): 521
 PEREZ Norma Beatriz (UNSL): 881, 1066
 PÉREZ Ramiro (UNLPAM): 1146
 PÉREZ Ricardo Daniel (UNJu): 238
 PÉREZ Santiago (UNdeC): 99
 PESADO Patricia (UNLP): 790, 588, 593, 599, 1141, 389
 PETRELLA Betina (UNaM): 760
 PETRIS Raquel H. (UNNE): 1217, 484
 PI PUIG Martin (UNLP): 935, 947
 PICCO Trinidad (UNNOBA): 1207
 PICCOLI María Fabiana (UNSL): 897, 958
 PICUCCI Marcos (UNCOMA): 465
 PIERGALINI Rosana (UNNOBA): 1207
 PINCIROLI Fernando: 706, 511
 PINTO LUFT Cristian (UNNE): 563
 PIOTROSKI Facundo Martin (UNaM): 505
 PIPINO Hugo (UTN-FRSFco): 978
 PIRAGINE María F. (UNNE): 484
 PIRRO Adriana (UNMdP): 1083
 PISARELLO María Inés (UNNE): 401
 PIZARRO Rubén (UNLPAM): 1146, 604
 POETA Tomás (UB): 376
 POLLO-CATTANEO María Florencia (UTN-FRBA): 1151
 POMPEI Sabrina (UNNOBA): 60
 PORTUGAL MURCIA Emmanuel Alejandro (UNdeC): 494, 869
 POUSA Adrián (UNLP): 1275, 941, 947, 935
 PREISEGGER Santiago (UNLP): 593
 PRIANO Rodolfo (UNLA): 446
 PRIMORAC Carlos R. (UNNE): 1078, 195
 PRINSCICH Nahuel (UNLP): 389
 PRINTISTA Marcela (UNSL): 345, 958
 PRISCHING Guillermo (UNTDF): 1192
 PUENTE Cristina (COMILLAS): 350
 PUIG Martin Pi (UNLP): 974
 PYTEL Pablo (UTN-FRBA): 1151

Q

QUEIRUGA Claudia (UNLP): 475, 489
 QUEREL Matías (UNLaM): 794
 QUINTEROS Sergio (UTN-FRC): 340
 QUIROGA Facundo (UNLP): 42
 QUIROGA Gabriel (UTN-FRM): 99
 QUIROGA Macarena (UNPSJB): 1106
 QUISPE Jairo Joel Maximiliano (UNJu): 238

R

RADELJAK Federico (UNLu): 214
 RAFFO Daniela (UB): 645
 RAMBO Alice (UNaM): 265,505
 RAMÓN Hugo (UNCPBA): 60, 992
 RAMÓN Hugo (UNNOBA): 997
 RAMOS Lautaro (UTN-FRCU): 325
 RAMOS Pablo Nicolás (UNJu): 300
 RAPALLINI José (UTN FRLP): 982
 RAPESTA Natalia (UTN-FRCU): 325
 RASJIDO José (UNPA): 50, 56
 RAUBER F. (UNaM): 265
 REBRIJ Romina (HIBA): 200
 REINOSO Elizabeth (UNJu): 568
 RETAMAR Soledad (UTN-FRCU): 325

REXACHS Dolores (UAB): 953
 REY M. (UNaM): 265
 REYES Nora (UNSL): 335, 345
 RIBA Alberto Eduardo (UNdeC): 770, 1171, 494, 869
 RIBEIRO Federico (UNLA): 441
 RICO Carlos (UNMdP): 1088
 RIESCO Daniel (UNSL): 1016, 660, 1051, 737
 RÍOS Domingo A. (UNNE): 109
 RIOS Leopoldo (UNNE): 859
 RIPODAS Alejandra (UNLP): 599
 RISSOLA Esteban A (UNLu): 214
 RIVERA Ramiro (UTN-FRCU): 325
 RIVERO David Sánchez (UNJu): 568
 RIVOIRA Alejandro Ramón (UNSL): 547
 ROBLEDO Ezequiel (UNO): 645
 ROCABADO Sergio (UNSa): 680
 ROCÍO Rodríguez (UAI): 748
 ROCÍO Rodríguez (UNLaM): 752
 RODRIGUEZ Anahí (UNLP): 614
 RODRÍGUEZ Carlos (UNLaM): 843
 RODRÍGUEZ Cintia Silvana (UNJu): 238
 RODRÍGUEZ Herlein Diego R. (UNPA): 159
 RODRIGUEZ Dario (UNLA): 446, 690
 RODRÍGUEZ EG Sebastián (UNLP): 935, 947
 RODRIGUEZ G. (UNSL): 396
 RODRÍGUEZ Ismael Pablo (UNLP): 947
 RODRIGUEZ Jorge (UNCOMA): 456, 465, 470, 500
 RODRIGUEZ Juan Manuel (UNTDF): 1186
 RODRIGUEZ Mariela (UNJu): 300
 RODRIGUEZ Marina (UNNOBA): 1207
 RODRIGUEZ Nelson R. (UNSJ): 897, 706
 RODRIGUEZ Omar (UTN FRLP): 982
 RODRÍGUEZ Rocío Andrea (UAI): 756, 1093, 526
 RODRÍGUEZ Sergio Elias (UNJu): 12
 ROGER Sandra (UNCOMA): 891
 ROGGERO Patricia (UNSL): 345
 ROJAS DELGADIN Cristian Alejandro (UNLPAM):
 1146
 ROJAS Sergio (UNTDF): 1186
 ROLDAN Janina (UNLPAM): 1222
 ROMAGNANO María (UNSJ): 205
 ROMANO Lucas (UNTDF): 848, 1186
 ROMERO Fernando (UNLP): 974
 ROMERO María Soledad (UTN-FRC): 340
 RONCHETTI Franco (UNLP): 42, 1296
 ROSADO Magdalena (UNLP): 389
 ROSALES Vanessa Aybar (UNLP): 489
 ROSAS Carolina (UGD): 290
 ROSAS M. V. (UNSL): 1166
 ROSATTO David (UNAJ): 912
 ROSENSTEIN Javier (UCH): 776, 897
 ROSENZVAIG Federico (UNSE): 675
 ROSETE-SUÁREZ Alejandro (ISPJAE): 42
 ROSSET Ana Lorena (UNC): 27
 ROTTOLI Giovanni (UTN-FRCU): 609
 ROVERO Mara Lia Inet (UNdeC): 670
 ROZAS Claudia (UNdeC): 953
 RUANO Darío (UNSL): 316
 RUCCI Enzo (UNLP): 947, 941
 RUEDA José (UNSJ): 635
 RUIDÍAS Héctor J. (UGD): 290
 RUIZ Susana Beatriz (UNSJ): 205
 RUSSO Claudia (UNCPBA): 60, 992, 1207
 RUSSO Claudia (UNNOBA): 997
 RUZ Cacilia Ana (CAECE): 285
 RYCKEBOER Hugo Emilio (UNLaM): 255

S

SAAD Gonzalo (UTN-FRC): 521
 SAAVEDRA Emmanuel (UNSE): 675
 SABOLANSKY Alejandro (UNLP): 1056
 SALAZAR MESÍ Natalí (UNLP): 1135
 SALDAÑO Viviana E. (UNPA): 558
 SALDARINI Javier (UTN-FRSFco): 542, 655
 SALDIVIA Alvaro (UNRN): 1212
 SALGADO Carlos (UNSL): 537, 542, 547, 270, 655,
 660, 685
 SALTO Carolina (UNLPAM): 70
 SALVATORE Juan E. (UNAJ): 153
 SAMBRANA Iván (UNNE): 563
 SAMEC Gustavo (UNPSJB): 229
 SAN MARTÍN Patricia (UNR): 451
 SANCHEZ Alberto (UNSL): 537, 542, 547, 270, 655,
 685
 SANCHEZ Carolina (UNLaM): 425
 SANCHEZ Mariano (UNLP): 1135, 941
 SANCHEZ PICCARDI Laura (UCSE): 275
 SANCHEZ ZUAIN Silvia (UCSE): 680
 SANDOVAL Iván Leandro (UNJu): 238
 SANTA MARÍA Cristobal R. (UNLaM): 200
 SANTANA Santiago (UNPSJB): 1106
 SANTANA Sonia (UNER): 1011, 1038
 SANTIAGO Martin (UNCPBA): 1130
 SANTOS D. (UNLA): 835

SANTOS Graciela (UNCPBA): 1197
 SANTOS Walter Benjamín (UNJu): 17
 SANZ Cecilia (UNLP): 1135, 1141, 1197, 1203, 1212, 389
 SANZ Victoria (UNLP): 941
 SAROBE Monica C. (UNNOBA): 992, 997
 SAVORGNANO Marcela (UTN-FRC): 1115
 SAYAGO Juan (UTN-FRM): 174
 SCANDOLO Ivan (UTN-FRC): 1115
 SCAPPINI Reinaldo (UTN-FRRe): 164, 104
 SCARELLO Eduardo (UTN-FRSFco): 716
 SCHAB Esteban (UTN-FRCU): 609, 325
 SCHARFF Lucia (UTN-FRSFco): 726
 SCHELSE Solange (UNaM): 417
 SCHIAVONI Alejandra (UNLP): 829, 1240
 SCHIAVONI Antonela (UNNOBA): 997
 SCHLAPS Erica (UNLPAM): 1222
 SCUTARI Paola Caymes (UTN-FRM): 922
 SEBASTIÁN Alberto (UB): 645
 SEGOVIA Dario (UNLPAM): 604, 420
 SELZER Matias Nicolas (UNS): 366
 SERAFINO Sandra (UNNOBA): 992
 SERÓN Natalia (UNPA): 56
 SERRA Silvio (UTN-FRC): 521
 SERRANO Diego J. (UTN-FRC): 1061
 SERRANO Eliana (UNNOBA): 1207
 SEVILLANO Gerardo (UTN-FRM): 838
 SILLERO Javier (UNSJ): 897
 SILVA Telmo (OTRA): 389
 SILVERA J. (UNSa): 710
 SILVESTRI Mario (UNSL): 480
 SKRIE Daniel (UK): 142
 SLAWISKI Javier (UNLaM): 843
 SMAIL Ana (UNNOBA): 1207
 SOLDÁN Carlos Paz (UNLu): 321
 SOLOZA Federico (UNCOMA): 799
 SOMMER Sonia (UNCOMA): 470
 SORIA Marcelo (UBA): 200
 SOSA BRUCHMAN Eugenia C. (UNCa): 7, 435
 SOSA Eduardo O (UGD): 123, 531, 184
 SOSA Juan Fernández (UNLP): 588
 SOSA Marcelo Omar (UNCa): 435
 SOSA Paula (UTN-FRC): 1061
 SOZZI Bianca (UTN-FRM): 430
 SPADONI Gustavo Fernando (UNJu): 300
 SPINELLI Adolfo (UNMdP): 1088
 SPOSITTO Osvaldo Mario (UNLaM): 255

STOPELLO Marta (UNNE): 484, 1217
 STORNI Antonio (UTN-FRM): 371
 STRACCIA Luciano (UTN-FRBA): 1151
 SUAREZ Gabriela irene (UNSE): 1203
 SUÁREZ Paula (UNRN): 1212
 SUELDO R. (UNaM): 505
 SUPPI Remo (UAB): 953
 SYNIUK Ruben (UGD): 184

T

TAFFERNABERR Carlos (UTN-FRM): 174
 TALAY Carlos A. (UNPA): 159
 TAMAGUSKU Martin (UNLPAM): 70
 TAMASHIRO Santiago (UNLaM): 1093
 TAPIA Marcela Alejandra (UNJu): 300
 TAPIA Maria Antonia (UNJu): 12, 17
 TAPIA Nestor (UNDEF): 1026, 1029
 TARANILLA María Teresa (UNSL): 280
 TARDIVO Laura (UNRC): 907, 922
 TARULLA Francisco (UB): 376
 TEJADA Jorge Damián (UNdeC): 494, 869
 TEJERO Germán (UNTDF): 1192
 TERAGNI Matías (UAI): 927, 1236
 TERZANO Julieta (UNNOBA): 997
 TESONE Fernando (UNLP): 588, 599
 TESSORE Juan Pablo (UNNOBA): 60
 TESTA Oscar (UNLPAM): 1146, 604
 TESTA Sylvia (CAECE): 2, 251
 TEXIER Jose (UNdeC): 770, 1171
 TEZE Juan Carlos Lionel (UNER): 65, 1306
 THOMAS Pablo J (UNLP): 790, 588, 599
 TIDONA Fernando (UNPSJB): 987
 TINETTI Fernando G. (UNLP): 974, 935
 TISERA Cristian (UNSL): 902
 TITIONIK Diamela (UNLPAM): 1222
 TITIOSKY Rolando (UB): 731
 TOBAR Sebastián (UTN-FRM): 174
 TOLABA Ana (UNJu): 238
 TOLOSA Gabriel (UNLu): 214, 220, 321
 TOLOSA Rodrigo (UNPSJB): 987
 TORALES Rodrigo (UNLP): 886
 TORRES Enrique Alba (UMA): 46
 TOSETTO Rita (UNdeC): 770
 TOURNOUD Adrián (UTN-FRCU): 609
 TRABES Guillermo (UNSL): 902
 TRASMONTANA Julio (UTN-FRSFco): 542, 655
 TREJO Stefano (UCSE): 680

TRIGILA Mariano (UCA): 930
TRIGUEROS Artemisa (UNLaM): 752
TRINIDAD Franco A. (UNPA): 159
TRIPPEL NAGEL Juan Manuel (UNS): 366
TROGLIA Carlos (UTN-FRM): 819
TROIANI Fernando Sanz (UNLPAM): 70
TUGNARELLI Monica (UNER): 1011, 1038

U

UNGARO Ana (UNLP): 475
URBANO Bárbara (UTN-FRC): 340
URIBE Federico (UNCOMA): 953
URQUIJO Ruben Ricardo (UNaM): 128, 505
URRIBARRI Dana K. (UNS): 361, 406
URRIZA José M. (UNPSJB): 987
USEGLIO Gustavo (UNNOBA): 992

V

VACA Pablo A. (UTN-FRC): 270
VALDEZ Jorge Ceferino (UNPA): 50
VALDEZ Nestor Jorge (UNCa): 7
VALLEJOS Oscar Adolfo (UCP): 189
VALLES Federico (UNLaM): 1093
VALIENTE Waldo (UNLaM): 854
VAN DE GRAAF Jeroen (UFMG): 1003
VARAS Valeria (UNPA): 50, 56
VARELA Pablo (UM): 179
VARGAS Alejandro (UNJu): 12,568
VARGAS Fernando (UNPA): 558
VARGAS Laura (UNC): 1061
VARGAS Luis Alejandro (UNJu): 300
VAZQUEZ Alejandro (UTN-FRM): 819
VAZQUEZ Juan Carlos (UTN-FRC): 84
VEGA Raúl Marcelo (UNCa): 7, 435
VEGLIA Jorge E. (UNNE): 104
VELA Florencia (UB): 376
VELAZCO Melisa (UTN-FRC): 89
VENOSA Paula (UNLP): 489, 1056
VERA Andrea (UNLaM): 573
VERA Ariel (UNO): 645
VERA BATISTA Fernando (UNDEF): 1047
VERA Cristina (UNSJ): 225
VERA LACEIRAS María Silvia (UNaM): 401
VERA María I (UNNE): 1217
VERA Pablo (UNLaM): 752, 526,1093
VERA PABLO Martin (UAI): 748,756

VERDEJO Ariel (UTN-FRM): 174
VERINO Claudia (UTN-FRSFco): 665
VIAÑA Graciela (UNSE): 824
VICENTE Néstor (OTRA): 953
VIDELA Lucas (UNLaM): 425
VIER Sergio Fabian (OTRA): 553
VIERA Leonel (UNTDF): 848
VILLCA Raúl (UNLaM): 854
VILCHES Diego (UNLP): 864
VILLA MONTE Augusto (UNLP): 350
VILLAFANE Juan Pablo (UNJu): 17
VILLAGRA Andrea (UNPA): 46, 50,56
VILLAGRA Silvia (UNPA): 558
VINCENZI Marcelo De (UAI): 1124
VIOLINI Lucía (UNLP): 1135
VITTORIO Marcos (UNLaM): 854
VOGEL Facundo (UNLaM): 1093
VOLKER Mariano (UNLaM): 854

W

WAGNER Laura (UNLPAM): 1222
WANZA Horacio Villagarcía (UNLP): 974, 935
WILLGING Pedro (UNLPAM): 420
WURN Guillermo (UNaM): 128

Y

YAMEL Laura (UNNOBA): 1207
YUAN Rebeca (UTN-FRSFco): 22

Z

ZABALA Gonzalo (UAI): 927, 1236
ZABALJAUREGUI Marcelo (UTN FRLP): 982
ZACCARDI Gonzalo (UP): 917
ZAIN Gaspar Acevedo (UNLaM): 526
ZAMBRANO Jusmeidy (UNdeC): 770
ZAMUDIO Eduardo (UNaM): 1286,265, 169
ZANARINI Dante (UNR): 451
ZANGARA Maria Alejandra (UNLP): 1135, 1186
ZAPANA José V (UNJu): 568
ZAPIRAIN Esteban (UNMdP): 1088,1098
ZARADNIK Ignacio J (UNLaM): 843
ZARATE Marcos (CENPAT): 229
ZEA CARDENAS Milagros (UTN-FRC): 1033
ZORATTO Valeria (UNCOMA): 624
ZORZAN Fabio (UNRC): 629

ZÚÑIGA Mariela Elizabeth (UNSL): 1166
ZURITA Rafael (UNCOMA): 456, 465

Agentes y Sistemas Inteligentes

Desarrollo de Habilidades sociales y cognitivas a través de Sistemas Inteligentes

Stella Maris Massa¹, Agustin Casamayor², Sylvia Testa³

Universidad CAECE/ Departamento de Sistemas/ Argentina
(7600) Olavarría 2464, +54-4993400

smassa4@gmail.com¹, acasamayor@ucaecemdp.edu.ar², stesta@caece.edu.ar³

RESUMEN

Este artículo presenta un proyecto de investigación bianual (2018-2020).

El presente proyecto de investigación se basa en la utilización de herramientas computacionales como Aprendizaje de Máquina (Machine Learning) e interfaces visuales adaptativas, para evaluar y asistir en el desarrollo de habilidades sociales y cognitivas.

Este sistema utilizará técnicas de Machine Learning para aprender de las interacciones de los participantes, a fin de poder crear un modelo que permita detectar posibles problemas sociales o de aprendizaje, para posteriormente asistir en su entrenamiento o compensación.

El uso de un entorno virtual que recree situaciones sociales cotidianas podría permitir evaluar en tiempo real los procesos de cognición social, poniendo en juego los indicadores más relevantes a través de experiencias directas con situaciones sociales que les permitan a las personas dar respuestas espontáneas.

Uno de los principales objetivos es que los participantes no se vean condicionados por los instrumentos de medición convencionales, pudiendo interactuar libremente en este entorno virtual, sin el sesgo de saber que están siendo evaluados. Los resultados se validarán a través de la comparación con aquellos obtenidos por los métodos tradicionales.

Palabras clave: Sistemas Inteligentes, Aprendizaje de Máquina, Teoría de la Mente, Aprendizaje.

CONTEXTO

El proyecto I+D+T que se presenta está radicado en el Departamento de Sistemas de la Universidad CAECE, Subsede Mar del Plata y Sede Buenos Aires, donde se dictan las Carreras Licenciatura e Ingeniería en Sistemas y el ante proyecto de Maestría en Ciencias de Datos e Innovación Empresarial. Desde el punto de vista de los recursos humanos, este proyecto contribuirá a mejorar la formación en temas de Inteligencia Artificial, Aprendizaje de Máquina, Big Data.

1. INTRODUCCIÓN

Para interactuar efectivamente en el ámbito social es necesario predecir las acciones de las personas. Uno de los procesos cognitivos que sirven a un buen funcionamiento social lo constituye la Teoría de la Mente (TdM).

El término “Teoría de la Mente” se refiere a una habilidad cognitiva compleja, que permite que un individuo atribuya estados mentales a sí mismo y a otros. Riviere y Núñez [1] consideran a la TdM como la competencia de atribuir mente a sí mismo y a otros, y de predecir y comprender sus conductas en función de entidades mentales como las creencias, deseos e intenciones. Es un subsistema cognitivo que se compone de un soporte conceptual y de mecanismos de inferencia, desempeñando la función de manejar, predecir e interpretar la conducta de otros [2].

La TdM contribuye a explicar el comportamiento de los demás, al mismo tiempo que permite desarrollar y adquirir una experticia social, asume procesos inobservables (deseos, creencias, emociones e

intenciones), que relacionados sistemáticamente explican y predicen como podría accionar una persona en circunstancias particulares. Este proceso se produce en presencia de otras personas, por lo cual se diferencia de la mera predicción de acontecimientos de orden físico. Mientras que las personas tienen experiencias internas, razones, motivaciones o intenciones de actuar, el comportamiento de las cosas físicas se explica completamente por las leyes de la física [3].

Las investigaciones en psicología evolutiva y en neurociencia social han proporcionado evidencias de que la mente humana posee capacidades especiales para procesar y adaptarse a situaciones complejas que emergen del entorno social [4]. Cabe considerar que la capacidad para atribuir intenciones a los demás es universal y su desarrollo ontogénico es muy homogéneo. Algunos modelos explicativos consideran que la TdM constituye una habilidad ya presente de manera incipiente desde el nacimiento y que progresa con el desarrollo, repitiendo los mismos pasos en los primeros cinco años aún en culturas muy diferentes [5]. Existe consenso en considerar la adquisición de la TdM entre los tres y los cinco años, aumentando su complejidad con la edad [6].

Los estudios empíricos de la última década se centraron en los procesos de atribución de estados mentales a otros, desde técnicas clásicas offline (fuera de una situación social concreta), donde se indagan aspectos que permiten comprender alguno de los componentes de la TdM de manera aislada, no valorando que sucede con dicho proceso online (tiempo real) [3]. Si bien autores como Slaughter y Repacholi [7] y Shamay-Tsoory [8] consideran que la TdM no es un constructo unidimensional, razón por la cual una prueba aislada no puede dar cuenta de su complejidad y progresión; las pruebas clásicas de TdM evalúan generalmente aspectos todo/nada sin atender a fases intermedias que pueden expresarse en un continuo evolutivo [9].

En nuestro medio, los estudios destinados a explorar el desarrollo progresivo de los

indicadores de TdM son limitados. Las restricciones principales se refieren, en primer lugar, a la ausencia de estudios que aborden de manera conjunta y exhaustiva el desarrollo continuo de la TdM en niños de entre 6 y 15 años y, en segundo lugar, a la ausencia de técnicas capaces de evaluar de manera online la TdM.

En este contexto, el campo de la Inteligencia Artificial (IA) y particularmente el Aprendizaje de Máquina puede aportar muchas soluciones a las necesidades que el ámbito de la TdM plantea. Ejemplos de estas tecnologías basadas en IA son los tutores inteligentes, los sistemas de gestión del aprendizaje, y los videojuegos [10].

En los últimos años la neurociencia social se ha beneficiado con el uso estímulos dinámicos de personajes virtuales animados [11]. Numerosas investigaciones han arrojado evidencia consistente de que no sólo la experiencia sino también las reacciones sociales, son sorprendentemente equivalentes en los encuentros sociales con mediadores virtuales en comparación con las interacciones directas cara a cara entre los seres humanos [12].

El uso de un entorno virtual que recree situaciones sociales cotidianas podría permitir evaluar de manera online los procesos de cognición social (específicamente la TdM), poniendo en juego los indicadores más relevantes a través de experiencias directas con situaciones sociales que les permitan a las personas dar respuestas espontáneas.

Según Vodeley [3] la creación de personajes virtuales, que pueden servir como seres humanos artificiales creíbles, proporciona una herramienta de investigación única para mejorar el conocimiento sobre los procesos psicológicos subyacentes y sus mecanismos neuronales.

Félix Eroles, director de proyectos en RedVisible.com y autor en el blog "Personas que Aprenden", aseguró en una entrevista con America Learning Media realizada en el año 2016, que una de las tecnologías que está explotando y que estará presente con más profusión es el Aprendizaje de Máquina. Sostiene que el aprendizaje estará guiado o

conducido por el Adaptive Learning, y que con ayuda de la IA, “podemos explorar las capacidades de los alumnos y determinar los ritmos, estilos de aprendizaje, estados emocionales, contenidos (...) que necesitan en el momento oportuno”.

En esta línea y en referencia a las tendencias actuales en materia de adaptive e-learning, José Omedes, Research & Development Director de iTopTraining, indicó en una entrevista al sitio “America Learning Media” en el año 2016, que en el aprendizaje basado en recomendación el software utiliza técnicas de Aprendizaje de Máquina para descubrir, de modo automático, los diferentes perfiles de usuario y sus itinerarios de aprendizaje óptimos a través de los contenidos, de modo que el sistema se realimente no sólo de la experiencia individual de un alumno en un curso, sino de la totalidad de los procesos de aprendizaje que han tenido lugar en la plataforma.

Otro ejemplo en el área de capacitación es la de la Fundación Inria Chile, una fundación sin fines de lucro creada por la institución pública francesa Inria (reconocida a nivel internacional por su aporte en I+D aplicada en el mundo digital, desde las Ciencias de la Computación y las Matemáticas Aplicadas). Dicha fundación trabaja conjuntamente con la Carnegie Mellon University (Pittsburg, Estados Unidos) en un equipo mixto llamado Eduband, que investiga cómo el Aprendizaje de Máquina puede generar y mejorar nuevas herramientas educativas. Concretamente, registra las reacciones de un usuario que está tomando, por ejemplo, un curso de capacitación, y desarrolla un algoritmo que analiza todos los datos y formula recomendaciones para facilitar el proceso de aprendizaje. La plataforma automáticamente realizará recomendaciones en temas que el estudiante aún no ha entendido y tratará con mayor rapidez los puntos que el alumno ya domina.

2. LINEAS DE INVESTIGACIÓN

Las líneas de investigación que se abordan en este proyecto son:

- Teoría de Mente
- Aprendizaje de Máquina Técnicas de Inteligencia Artificial

3. RESULTADOS ESPERADOS Y OBJETIVOS

Se plantea como Objetivo General: “Generar un sistema inteligente con interfaces visuales adaptativas, para evaluar y asistir en el desarrollo de habilidades sociales y cognitivas”.

El presente proyecto de investigación se basa en la utilización de herramientas computacionales como Aprendizaje de Máquina (Machine Learning) e interfaces visuales adaptativas, para evaluar y asistir en el desarrollo de habilidades sociales y cognitivas.

Este sistema utilizará técnicas de Machine Learning para aprender de las interacciones de los participantes, a fin de poder crear un modelo que permita detectar posibles problemas sociales o de aprendizaje, para posteriormente asistir en su entrenamiento o compensación. Uno de los principales objetivos es que los participantes no se vean condicionados por los instrumentos de medición convencionales, pudiendo interactuar libremente con el videojuego, sin el sesgo de saber que están siendo evaluados. Asimismo, este tipo de evaluación puede escalar de manera sencilla, a fin de poder obtener un mayor número de muestras en menor tiempo y requiriendo una cantidad menor de evaluadores humanos.

En particular, se proponen los siguientes objetivos específicos:

- Explorar y analizar los principios básicos de diseño de sistemas inteligentes con interfaces visuales adaptativas.
- Explorar y analizar los fundamentos computacionales, matemáticos y estadísticos del Aprendizaje de Máquina aplicado al diagnóstico y adquisición de habilidades sociales y cognitivas.
- Seleccionar las tecnologías que permiten llevar a cabo el desarrollo de un sistema inteligente con interfaces visuales adaptativas para evaluar y asistir en el desarrollo de habilidades sociales y cognitivas.
- Formular un modelo que permita detectar posibles problemas sociales o de aprendizaje, para posteriormente asistir en su entrenamiento o compensación.
- Construir una aplicación interactiva inmersiva basada en el modelo formulado.
- Realizar pruebas piloto de implementación.

Como avances del proyecto, se realizará un Convenio de colaboración entre el Instituto de Psicología Básica, Aplicada y Tecnología de la Facultad de Psicología de la Universidad Nacional de Mar del Plata (IPSIBAT) y el Departamento de Sistemas de la Universidad CAECE.

El presente convenio tiene por objeto la cooperación en el marco del desarrollo del proyecto “Procesamiento cognitivo y afectivo de la teoría de la mente y sus relaciones con el funcionamiento social en niños y adolescentes”, con aportes de instrumentos relativos a Sistemas de Aprendizaje de Máquina, entre el Grupo Comportamiento Humano, Genética y Ambiente de la UNMDP, perteneciente al IPSIBAT, y el Proyecto de Investigación “Desarrollo de Habilidades sociales y cognitivas a través de Sistemas Inteligentes”.

El objetivo del convenio es llevar a cabo trabajos de investigación, transferencia tecnológica y capacitación en el campo de las Neurociencias Sociales aplicando conceptos de Sistemas Inteligentes.

4. FORMACIÓN DE RECURSOS HUMANOS

Se acreditan dos tesis doctorales: Doctorado en Ciencias Informáticas de la Facultad de Informática de la UNLP y Doctorado en Ciencias de la Computación de UNICEN y una Maestría en Tecnología de la Información de CAECE.

Participan en el proyecto alumnos avanzados de las carreras Ingeniería y Licenciatura en Sistemas de la Universidad CAECE, sede Mar del Plata.

5. REFERENCIAS BIBLIOGRÁFICAS

- [1] Rivière, A. & Nuñez, M. (1996). *La mirada Mental: desarrollo de las capacidades cognitivas interpersonales*. Buenos Aires: Aique.
- [2] Tirapu, J., Pérez, G., Erekatxo, M. & Pelegrín, C. (2007). ¿Qué es la teoría de la mente?. *Revista de Neurología*, 44(8), 479-489.
- [3] Vogeley, K. (2017). Two social brains: neural mechanisms of intersubjectivity. *Phil. Trans. R. Soc. B*, 372(1727), 20160245.
- [4] Vogeley, K., & Bente, G. (2010). “Artificial humans”: Psychology and neuroscience perspectives on embodiment and nonverbal communication. *Neural Networks*, 23(8), 1077-1090.
- [5] Wellman, H., Cross, D. & Watson, J. (2001). Meta-analysis of theory of mind development. The truth about false beliefs. *Child development*, 72 (3), 655-684.
- [6] Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The reading the mind in the eyes test revised version: a study with normal adults and adults with Asperger syndrome or high-

- functioning autism. *J Child Psychol Psychiatry*, 42, 241-51.
- [7] Slaughter, V., y Repacholi, B. (2003). Individual differences in theory of mind: what are we investigating? En B. Repacholi y V. Slaughter (Eds.) *Individual differences in theory of mind: Implications for typical and atypical development*, (pp. 1-12). Hove, UK: Psychology Press.
- [8] Shamay-Tsoory, S., Harari, H., Aharon-Peretz, J., & Levkovitz, L. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies, *Cortex*, 46, 668-677.
- [9] Valdez, D. (2007). *Necesidades educativas especiales en trastornos del desarrollo*. Buenos Aires: AIQUE.
- [10] Silverstein, S. (2006). *Colleges see the future in technology*. Los Angeles Times.
- [11] Adolphs, R. (2003). Cognitive neuroscience of human social behavior. *Nature Reviews Neuroscience*, 4(3), 165–178.
- [12] Garau, M., Slater, M., Pertaub, D. P., & Razaque, S. (2005). The responses of people to virtual humans in an immersive virtual environment. *Presence*, 14(1), 104–116.

Operadores de Cambio de Creencias No-Priorizada en Sistemas Argumentativos Aplicado para Programación Lógica Rebatible

Néstor Jorge Valdez[†] Eugenia Cecilia Sosa Bruchmann[†]
Raúl Marcelo Vega[†]

[†] Departamento de Ciencias de la Computación, Fac. de Ciencias Exactas y Naturales
Universidad Nacional de Catamarca (UNCa)
Av. Belgrano 300 - San Fernando del Valle de Catamarca
Tel.: (03834)420900
e-mails: njvaldez@exactas.unca.edu.ar
sosab-ec@hotmail.com
rmvega68@yahoo.com.ar

Resumen

Un sistema de cambio de creencias constituye un marco lógico para la modelización de la dinámica de conocimiento. Es decir, cómo modificamos nuestras creencias cuando recibimos nueva información. En revisión de creencias, incorporamos nueva información y descartamos cierta información existente para acomodar la recién ingresada. Sin embargo, en recientes investigaciones algunos autores argumentan que la nueva información no siempre debería tener *prioridad* sobre la información existente, esto llevo a que se propusieran algunos métodos de revisión de creencias en los que *no se acepta* necesariamente la nueva información. Por ello, es que resulta de interés estudiar las operaciones que se denominan de *cambio no-priorizada*. Esta línea de investigación tiene como objetivo el estudio de nuevos tipos de operadores no-priorizada potenciando su intuición de procesamiento con operadores basados en *explicaciones* (sistemas de argumentación). Se estudiarán los *operadores de*

revisión en particular en donde algunos de ellos procesan la información con el apoyo de una *explicación*. También, se establecerá procedimientos de aplicación de estas construcciones en *sistemas argumentativos* que brindarán soporte a las capacidades de razonamiento de agentes que proveen información que puede ser incompleta y/o contradictoria y de esta manera mejorar los mecanismos de toma de decisión automática. En tal sentido, se espera que los resultados obtenidos brinden una nueva perspectiva para desarrollar herramientas tecnológicas que funcionen desde un enfoque multi-agente.

Palabras Claves Cambio de Creencias No-Priorizada, Razonamiento Argumentativo, Programación Lógica Rebatible.

1. Contexto

Esta línea de investigación se realizará dentro del ámbito del Laboratorio de Investigación del Departamento de Ciencias de la Computación de la Facultad de Cien-

cias Exactas y Naturales de la Universidad Nacional de Catamarca. Esta asociado con el proyecto de investigación Anual: “Operadores de Revisión de Creencias Múltiples No-Priorizada en Sistemas Argumentativos Aplicado para Programación Lógica Rebatible”. Financiado por el programa de desarrollo científico y tecnológico de la Secretaría de Ciencia y Tecnología: Consejo de Investigación, de la Universidad Nacional de Catamarca. *Periodo* :01/01/2018 al 31/12/2018.

2. Introducción

Un sistema de revisión de creencias constituye un marco lógico para la modelización de la dinámica de conocimiento. Es decir, cómo modificamos nuestras creencias cuando recibimos nueva información [1]. La revisión de creencias ha sido ampliamente debatido en los últimos años. En revisión de creencias, se incorpora nueva información y se descarta cierta información existente para acomodar la recién ingresada.

Sin embargo, muchos investigadores argumentan que la información nueva *no siempre debería tener prioridad sobre la información existente* por lo que se han propuesto algunos métodos de revisión de creencias *no-priorizada* en los que no se acepta necesariamente la nueva información. Por ejemplo, la *semi-revisión* introducida por Hansson [8, 9], ella difiere de la revisión de creencia clásica en dos aspectos: primero, la información original esta representada como una base de creencias en lugar de un conjunto de creencias, y segundo, la nueva información no es siempre aceptada. Fuhrmann [6] propone una forma diferente de revisión no-priorizada. El define un *operador de mezcla* (merge operator) en el cual dos bases de conocimiento pueden combinarse en una. El operador de mezcla abre la posibilidad de

que la nueva información sea parcialmente o totalmente ignorada si la vieja información es más fuerte. Este tipo de operador permite dos caminos *no posibles* en los operadores de revisión AGM: aceptar parcialmente la nueva información o rechazarla totalmente. La operación de mezcla une la vieja información con la nueva, sin darle precedencia o prioridad a ninguna de ellas y elimina posibles contradicciones.

Los operadores de revisión no-priorizada vistos hasta ahora tienen una característica distintiva: o aceptan plenamente la sentencia a incorporar o la rechazan totalmente. Esto es, o son muy osados o muy cautos en su actitud epistémica. No aceptan una solución intermedia al problema de aceptación de una sentencia. Por tal motivo, Fermé y Hansson [5] proponen un operador de revisión que puede aceptar parte de la información que constituye la entrada epistémica. Por este motivo, es que los autores denominaron a este operador *revisión selectiva* (selective revision). El operador de revisión con *límite de credibilidad* es otro tipo de operador no-priorizada [10]. El mismo utiliza una construcción alternativa del operador de *screened revision* definido por Makinson [7]. Si debemos revisar un conjunto de creencias K con respecto a una sentencia α debemos recurrir a un conjunto C , denominado conjunto de creencias creíbles. Este conjunto, determina que sentencias (creencias o no-creencias) podrían ser aceptadas en un proceso de revisión. Por último, el operador de revisión mediante *explicaciones* es un nuevo tipo de operador no-priorizada [2, 3, 4]. El mismo es aplicable sobre bases de creencias y se basa en la siguiente intuición: antes de incorporar una nueva creencia a nuestro conocimiento, exigimos una explicación para la misma. Esta explicación se representa con un conjunto de sentencias con ciertas restricciones. Luego, confrontamos esa explicación con nues-

tras propias creencias (pertenecientes a la base de creencias). Si la explicación 'sobrevive' al debate, entonces la explicación de la nueva sentencia es aceptada. De lo contrario, la base de creencias original permanece sin cambios. La construcción de este operador se basa en operadores anteriormente presentados.

3. Línea de Investigación y Desarrollo

Esta línea de investigación toma como punto de partida los aportes en la temática de los operadores de contracción y revisión múltiples priorizada de la teoría de cambio de creencias bajo clausulas Horn que han sido ampliamente investigado por los autores de este artículo [11, 12, 13, 15, 14]. Uno de los objetivos principales es el estudio de las operaciones de cambio no-priorizada. En este contexto, consideramos dos tipos de cambios: el primero, los operadores de cambio no-priorizada en donde todas las nuevas creencias no siempre deben ser aceptadas (y entre ellas, nuestro principal interés se centra en los operadores con *credibilidad limitada*), y el segundo, se refiere a operadores de *mezcla* (merging) que permite que creencias antiguas y nuevas jueguen roles simétricos dentro de un proceso de cambio.

Así también, dentro de la teoría de cambio, existen diversos tipos de operadores de revisión en donde algunos de ellos procesan la información con el apoyo de una *explicación*. Un operador de revisión puede modificar ya sea el conocimiento rebatible o no rebatible. El problema principal es determinar si alguna pieza de información es no rebatible o rebatible. Una solución simple podría ser la incorporación de conocimiento directamente en el conocimiento rebatible. Pero esta solución es demasiado simple y no es muy realista. La calificación del conocimien-

to es dinámica, es decir, que evoluciona con el tiempo y en la incorporación de nueva información. Cuando un agente incorpora conocimiento por lo general incorpora conocimiento no rebatible. Es por ello, que en los sistemas argumentativos resulta necesario contar con criterios de preferencia entre argumentos. La implementación de este tipo de sistemas está dada por los servicios de razonamiento basados en Programación Lógica Rebatible.

Las Teorías de Cambio de Creencias y Argumentación (principales ejes de esta investigación) pertenecen al amplio campo de la Representación del Conocimiento, sin embargo, sus puntos focales básicos son diferentes. Ahora, las interrelaciones entre los dos campos están todavía en su mayor parte sin explorar. Tanto los campos de la teoría de la argumentación como la de revisión de creencias son de importancia sustancial para los sistemas multi-agente que se enfrentan a un uso intensivo en aplicaciones prácticas industriales. Sumado a ello, el creciente uso de métodos y herramientas de la teoría de cambio de creencias en la teoría de la argumentación y viceversa.

El presente proyecto, tiene por finalidad volcar resultados y mejoras que se obtengan en el desarrollo de modelos de revisión de creencias no priorizada combinado con argumentación y de esta manera agilizar las capacidades de razonamiento mediante la implementación de la programación lógica rebatible (DeLP).

4. Resultados y Objetivos

Diversas contribuciones relacionado a la temática de esta investigación, fueron presentados en el ámbito de las ciencias básicas, provocando un impacto directo en el desarrollo de áreas tecnológicas de vanguardia

en Ciencias de la Computación, tales como la especificación de instituciones virtuales, desarrollo y automatización de mercados, subastas electrónicas a través de Internet, mejoras de mecanismos de toma de decisión automática, etc.

En esta línea de investigación se espera obtener los siguientes resultados:

- Ampliar y combinar la utilidad de las construcciones con sentencias múltiples de tipo no-priorizada y simétricos (mezcla) a otros tipos de cambios como los basados en explicaciones.
- Investigar y determinar las interrelaciones entre dos campos de la representación de conocimiento como son la teoría de argumentación y cambio de creencias, considerando que son diferentes sus puntos focales, lo que promoverá el empleo de métodos basados en argumentación caracterizados por la adopción de elementos de la teoría de cambio de creencias.
- Formular procedimientos basados en programación lógica rebatible (DeLP), para mejorar las capacidades de razonamiento de agentes en operaciones de revisión ejecutados por sistemas argumentativos.

Respecto a los objetivos de esta investigación, se espera obtener contribuciones en el área de las ciencias básicas y en el ámbito de aplicaciones tecnológicas.

Por último, se integrará programas de capacitación e intercambio para los integrantes del proyecto y otras áreas afines. Estas actividades serán de carácter:

- técnica: estudio y manejo de herramientas informáticas como DeLP y demás programas existentes en el campo de estudio; y

- metodológica: estrategias de investigación científica, mecanismos de publicación de artículos científicos, planificación de tutorías en ambientes presenciales y virtuales.

5. Formación de Recursos Humanos

Dentro de la temática de esta línea de investigación: *los operadores de revisión no-priorizada combinado con operadores basados en explicaciones (sistemas argumentativos) de la teoría de cambio de creencias* han sido ampliamente investigado por parte de los autores de esta contribución, siendo uno de los temas de tesis para alcanzar el grado de Magister en Ciencias de la Computación en la UNS (Argentina) de Néstor Jorge Valdez (uno de los autores de este artículo).

Así también, se espera que durante el desarrollo del proyecto, los integrantes del mismo que estén en proceso de elaboración de su tesis de posgrado puedan consolidar su formación en investigación, y que el trabajo realizado contribuya a su graduación.

Referencias

- [1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50, 1985.
- [2] Marcelo A. Falappa and Guillermo R. Simari. Propiedades del operador de revisión mediante argumentos. *Primer Congreso Argentino de Ciencias de la Computación, CACIC 95*, pages 325–336, 1995.
- [3] Marcelo A. Falappa and Guillermo R. Simari. Construction of revisions by ex-

- planations. *Cuarto Congreso Argentino de Ciencias de la Computación, CACIC 98*, pages 933–946, 1998.
- [4] Marcelo A. Falappa and Guillermo R. Simari. Explanations, belief revision and defeasible reasoning. *Cuarto Congreso Argentino de Ciencias de la Computación, CACIC 98*, pages–, 1998.
- [5] Eduardo L. Fermé and Sven Ove Hansson. Selective revision. *Studia Logica*, 63:331–342, 1998.
- [6] Fuhrmann. An essay on contraction. studies in logic, language and information. *CSLI Publications, Stanford*, 1997.
- [7] David Makinson. Screened revision. *In Theoria of Hansson S. 1997*, pages–, 1997.
- [8] Hansson S. Semi-revision. *Journal of Applied Non-Classical Logic*, pages 151–175, 1997.
- [9] Hansson S. A survey of non-prioritized belief revision. *Erkenntnis*, 50:413–427, 1999.
- [10] John Cantwell Sven Ove Hansson, Eduardo Fermé and Marcelo Falappa. Credibility limited revision. pages–, 1998.
- [11] Valdez and Falappa. Dinámica de conocimiento: Contracción múltiple en lenguajes horn. *XIX Congreso Argentino de Ciencias de la Computación, XIV Workshop Agentes y Sistemas Inteligentes (WASI), CACiC'2013*, 2013.
- [12] Valdez and Falappa. Dinámica de conocimiento: Contracciones horn a partir de ordenamientos epistémicos. *42JAIIO Jornadas Argentinas de Informáticas, ASAI 2013, 42 JAIIO'2013*, pages 206–209, 2013.
- [13] Valdez and Falappa. Implementación para bases de creencias horn de operadores de contracción múltiple. *XX Congreso Argentino de Ciencias de la Computación, XV Workshop Agentes y Sistemas Inteligentes (WASI), CACiC'2014*, 2014.
- [14] Valdez and Falappa. Multiple revision on horn belief bases. *XXII Congreso Argentino de Ciencias de la Computación, XVII Workshop Agentes y Sistemas Inteligentes (WASI), CACiC'2016*, 2016.
- [15] Valdez, Lara, Pedraza, and Teseira. Dinámica de conocimiento: Cambio de creencias múltiples. temática de investigación de inteligencia artificial en las ciencias de la computación. *JUCEN'15, Facultad de Ciencias Exactas y Naturales, UNCa. Catamarca 9 y 10 de diciembre de 2015 - Argentina*, pages 30–36, 2015.

Aplicación de técnicas de Inteligencia Artificial para evaluar la producción científico-académica de investigadores de Universidades públicas del Noroeste Argentino.

José Federico Medrano, Alejandro Vargas, Miguel Augusto Azar, Fernando Rubén Aramayo, Sergio Elias Rodriguez, Maria Antonia Tapia, Alfredo Ariel Diaz

jfedericomedrano@gmail.com, aleva98@yahoo.com, auazar@live.com,
fernando.ruben.aramayo@gmail.com, egiors@gmail.com, maria.antonia.tapia@gmail.com,
alfredodiazjy@gmail.com

VRAIn / Visualización y Recuperación Avanzada de Información / Facultad de Ingeniería
Universidad Nacional de Jujuy - Ítalo Palanca 10, +54 (388) 4221587

RESUMEN

La evaluación de la producción científica o específicamente la evaluación de la productividad de un científico, ha sido desde que se iniciaran las publicaciones de los resultados de la investigación, una tarea nada sencilla. Los indicadores bibliométricos se encargan de contabilizar y representar en cierta medida cuan productivo es un investigador, contabilizando la cantidad de publicaciones y la cantidad de citas recibidas por estas. Pero para tener una imagen completa, es necesario agregar otro aspecto relacionado con el impacto y popularidad de estas publicaciones, para ello el uso de las redes sociales académicas enriquecen los indicadores bibliométricos dándoles un sentido mucho más amplio, por ejemplo: cantidad de veces que un contenido es descargado, comentado, compartido y leído entre tantos otros, alternativas *altmetrics*. Este proyecto plantea un estudio bibliométrico transversal a todas las Universidad públicas del Noroeste Argentino empleando herramientas y técnicas de recuperación y visualización de

información, dos ramas de la inteligencia artificial que han tenido un gran desarrollo en la última década. Contar con un panorama de estos aspectos no solo reflejará la presencia y popularidad de las investigaciones realizadas por estos científicos, sino también sentará las bases para poder comparar el estado de cada una de las universidades.

Palabras clave: *Bibliometría, producción científico-académica, redes sociales académicas, indicadores, altmetrics*

CONTEXTO

La línea de investigación aquí presentada se encuentra inserta en el proyecto: “*Aplicación de técnicas de Inteligencia Artificial para evaluar la producción científico-académica de investigadores de Universidades públicas del Noroeste Argentino*”, ejecutado a partir de enero de 2018 con una duración de 2 años. Dicho proyecto es llevado a cabo por el grupo de investigación VRAIn (Visualización y Recuperación Avanzada de Información) de la Facultad de Ingeniería de la Universidad

Nacional de Jujuy. El proyecto se encuentra acreditado y financiado parcialmente por la Secretaría de Ciencia y Técnica y Estudios Regionales de la Universidad Nacional de Jujuy (Resolución R. N° 2928/17)

1. INTRODUCCIÓN

El crecimiento de la producción científica en las últimas décadas así como su recopilación en bases de datos bibliográficas automatizadas ha potenciado el uso de la "Bibliometría" (Aström, 2007) (White & McCain, 1989) y la generación de indicadores para medir los resultados de la actividad científica y tecnológica. Los indicadores bibliométricos son datos estadísticos deducidos de las distintas características de las publicaciones científicas, en base al importante papel que desempeñan estas en la difusión y transmisión del conocimiento generado en la investigación. Son válidos cuando los resultados de la investigación se transmiten a través de publicaciones científicas y técnicas. Proporcionan información cuantitativa y objetiva sobre los resultados del proceso investigador, su volumen, evolución, visibilidad y estructura, pero no informan sobre los progresos del conocimiento (Mainardi & De Morán-Suárez, 2011).

Una investigación profunda y exhaustiva es llevada a cabo por (Waltman, 2016) sobre los indicadores de impacto de citas, tomando como fuente de datos *Web of Science (WoS)*, *Scopus* y *GS (Google Scholar)*. Este estudio analiza los indicadores más importantes relacionados con el recuento de citas de publicaciones, teniendo en cuenta las diferencias que existen al comparar estos indicadores entre campos distintos, indicando que para ello es necesario un proceso de normalización de datos, ya que reconoce que existen áreas donde la publicación de material es mucho más frecuente que en otras, por ejemplo, un bioquímico con 25 citas no puede considerarse que posee un impacto de citas mayor que un matemático con 10 citas. Lo mismo sucede con el año de publicación,

existen casos en que hay años más prolíficos que otros, y en el mismo sentido el tipo de publicación, sea que se trata de un artículo, una revisión o libro.

Hoy en día el análisis o recuento de citas bibliográficas es el mecanismo más efectivo o al menos el más utilizado para medir la productividad científico-académica, no solo por su sencillez de aplicación sino por su transparencia. Constituye un pilar a la hora de conceder subvenciones y financiamientos de proyectos y en la toma de decisiones, ya que son cada vez más utilizados como una medida de desempeño por el cual los científicos y profesores se clasifican, son promovidos y financiados, además el análisis de citas permite a los investigadores dar seguimiento al desarrollo e impacto de un artículo a través del tiempo (Medrano, 2017).

Según (Capaccioni & Spina, 2012), "La cuestión más debatida es, por supuesto, cuáles son los criterios de evaluación. ¿Criterios únicos o criterios diferentes para las diferentes áreas? ¿Criterios cualitativos o criterios cuantitativos? La importancia sobre la evaluación de la investigación científica se ha incrementado consistentemente en la primera década del nuevo siglo, sumado a la aplicación de algoritmos sofisticados que actualmente se encuentran disponibles para el análisis de citas permitiendo evaluar la "calidad" de estas, permitió la proliferación de un gran número de indicadores, entre ellos dos grupos bien diferenciados, los que se encargan de evaluar una revista científica y los que evalúan la productividad de un científico.

Además de las bases de datos tradicionales para evaluar la productividad científica y además de los distintos indicadores mencionados, existen otro tipo de iniciativas dirigidas a medir el impacto de los materiales publicados online, de forma más amplia que con el recuento de citas. Estas iniciativas son conocidas como *altmetrics* o métricas alternativas (Priem, Taraborelli, Groth, & Neylon, 2017). El término *altmetrics* fue introducido por primera vez en un Tweet en 2010 y más tarde (Priem, Altmetrics, 2014)

amplió la definición para incluir medidas de impacto académicas disponibles en cualquier plataforma en línea.

El conjunto de métricas comúnmente conocida como *altmetrics*, se basan generalmente en la medición de la actividad en línea relacionada con los académicos o los contenidos académicos derivados de las redes sociales y plataformas web 2.0 (Martín-Martín, Orduña-Malea, Ayllon, & López-Cózar, 2016). Sin embargo, la definición de lo que constituye un indicador *altmetrics* está en constante cambio, ya que está determinada en gran medida por las posibilidades técnicas y, más específicamente, la disponibilidad de Interfaces de Programación de Aplicaciones (API) (Haustein, Sugimoto, & Larivière, 2015).

La motivación de este proyecto resulta clara, ya que contar con mecanismo y herramientas para evaluar objetivamente los resultados de la investigación de un científico o grupo de investigación es importante.

Una muestra de la importancia de la evaluación científica son los rankings que se publican de forma anual para establecer el nivel de las universidades, y del mismo modo los rankings de los investigadores más prolíficos. Es más, estos indicadores tanto a nivel personal o institucional resultarían de utilidad al momento de conceder subvenciones y financiamientos de proyectos y en la toma de decisiones, ya que son cada vez más utilizados como una medida de desempeño por el cual los científicos y profesores se clasifican, son promovidos y financiados, ejemplo de ello es la evaluación a través de los sexenios de investigación de universidades europeas. Este instrumento, el sexenio investigador, es considerado en la mayor parte de los ámbitos científicos como una medida incuestionable de la calidad de la actividad investigadora, dado su elevado grado de aceptación por la comunidad universitaria (García-Berro, Amblàs, Sallarés, Bugada, & Roca, 2013).

Por esta razón, conocer el estado actual de cada universidad del NOA en cuanto a la cantidad de publicaciones y al impacto de estas, ofrecerá un

panorama adecuado además de marcar el camino para poder comparar cada institución con el resto de universidades del país.

En estos últimos años evaluar y cuantificar la productividad científica se ha convertido en tema de un profundo análisis ya que las cantidades de información aumentan constantemente y se vuelven cada vez más dinámicas (Miguel & Dimitri, 2013). En este momento existen variados métodos cuantitativos para medir esta producción, entre los más conocidos y usados destacan: el factor de impacto y *h-index* (Hirsch, 2005). Si bien estos números reflejan la relación entre artículos publicados y la popularidad de los mismos, entiéndase esto en términos de la cantidad de veces que un artículo es citado por otro o la importancia asignada a la revista en donde se publique, muy a menudo no resultan suficientes o precisos.

Los motores de búsqueda académicos de libre acceso como *Google Scholar* y *Microsoft Academic*, en contraposición a las bases de datos bibliográficas tradicionales como *Scopus* o *Web of Science*, presentan ciertas ventajas: El rango de cobertura es amplio, no solo artículos de revistas y congresos sino libros, tesis, informes y muchos otros tipos de documentos, son herramientas de libre acceso y gratuitas, no están circunscritos a determinadas áreas del conocimiento, y cada vez el alcance es mayor hacia áreas de menor relevancia, son muy populares y de fácil uso, son útiles para cualquier trabajo académico sin importar la envergadura de este.

Y para darle un mayor significado a estos indicadores, poder evaluar la presencia y la utilidad que se le da a estas publicaciones en redes sociales académicas, propone un enfoque desde otro ángulo. Hoy en día, ambos enfoques son distintos pero complementarios, valerse del uso de solo un tipo de indicadores no entregaría toda la información o solo se limitaría a evaluar un aspecto, por ello la necesidad de contar con los dos protagonistas de la historia, por un lado los indicadores bibliométricos tradicionales encargados de contabilizar la cantidad de

publicaciones y citas recibidas, y por otro lado los indicadores alométricos encargados de evaluar el impacto en redes sociales de dichas publicaciones.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación propone estudiar, evaluar y aplicar distintas técnicas de inteligencia artificial para evaluar la producción de los científicos de universidades públicas del NOA. Para ello se pretende utilizar métodos de recuperación de información y de visualización de información para conformar una base de datos tanto de los investigadores como de las publicaciones, realizar una limpieza de los mismos (empleando técnicas de detección de duplicados, fusión de citas) y poder obtener nuevo conocimiento a partir de ello, realizando comparaciones, inferencias y predicciones. En este sentido se emplearán algoritmos de estas áreas de la inteligencia artificial para lograr el propósito deseado.

3. RESULTADOS OBTENIDOS/ESPERADOS

El presente proyecto se centra en conocer el estado actual de la producción científico-académica de los investigadores de las universidades públicas del Noroeste Argentino

Particularmente se espera lograr:

- Analizar las técnicas de recuperación automática de información.
- Conformar una base de datos de publicaciones de investigadores libre de sesgos y normalizada.
- Conocer el estado actual y la presencia de investigadores en redes sociales académicas.
- Desarrollar y adaptar las técnicas de visualización de información para presentar los resultados obtenidos.
- Desarrollar una aplicación que permita procesar y evaluar los datos recolectados.

- Calcular indicadores bibliométricos y alométricos sobre los datos procesados.
- Realizar predicciones y obtener nuevo conocimiento a partir de los datos existentes.
- Evaluar el alcance de los resultados de la investigación, elaborar rankings y publicar las conclusiones alcanzadas.

4. FORMACIÓN DE RECURSOS HUMANOS

Durante la realización de este proyecto se espera lograr, como mínimo, la culminación de 1 tesis de maestría y 2 tesis de grado dirigidas y/o codirigidas por los integrantes del proyecto y relacionadas con la temática.

Una tesis doctoral, relacionada ampliamente con la línea de investigación presentada (Medrano, 2017), emplea varias técnicas y herramientas que se utilizarán en la implementación de este proyecto.

Finalmente, es constante la búsqueda hacia la consolidación como investigadores de los miembros más recientes del grupo.

5. BIBLIOGRAFÍA

- Aström, F. (2007). Changes in the LIS Research Front: Time-Sliced Cocitation Analyses of LIS Journal Articles, 1990-2004. *Journal of the American Society for Information Science and Technology*, 58(7), 947-957.
- Capaccioni, A., & Spina, G. (2012). Italian SSH journals in Journal Citation Reports (JCR) and in SCImago Journal Rank (SJR): data and first analysis. *Italian Journal of Library & Information Science*, 3(1), 4787-4807.
- García-Berro, E., Amblàs, G., Sallarés, J., Bugada, G., & Roca, S. (2013).

- Docencia e investigación: ¿un falso dilema? *Aula Abierta*, 41(2), 13-22.
- Haustein, S., Sugimoto, C. R., & Larivière, V. (2015). Guest editorial: Social media metrics in scholarly communication. *Aslib Journal of Information Management*, 67(3).
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 16569-16572.
- Mainardi, C. F., & De Morán-Suárez, M. A. (2011). La responsabilidad social corporativa (rsc) en las bases de datos scopus y wos (estudio bibliométrico). *EDICIC*, 1(4), 141-160.
- Martín-Martín, A., Orduña-Malea, E., Ayllon, J. M., & López-Cózar, E. D. (2016). The counting house: measuring those who count. presence of bibliometrics, scientometrics, informetrics, webometrics and altmetrics in the google scholar citations, researcherid, researchgate, mendeley & twitter. *CoRR*.
- Medrano, J. F. (2017). *Evaluación de la producción científica mediante motores de búsqueda académicos y de acceso libre*. Tesis doctoral, Universidad de Salamanca, Informática y Automática, Salamanca.
- Miguel, S., & Dimitri, P. J. (2013). La investigación en bibliometría en la Argentina: quiénes son y qué producen los autores argentinos que realizan estudios bibliométricos. (29), 117-138.
- Priem, J. (2014). Altmetrics. (B. Cronin, & C. R. Sugimoto, Edits.) *Beyond bibliometrics: harnessing multidimensional indicators of performance*, 263-287.
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (20 de 08 de 2017). *Altmetrics: A manifesto*, (v.1.0). Obtenido de <http://altmetrics.org/manifesto>
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Informetrics*, 10(2), 365-391.
- White, H., & McCain, K. W. (1989). Bibliometrics. *Annual Review of Information Science and Technology*, 24, 119-186.

Diseño de Algoritmos Inteligentes Aplicados a Interfaces Humano-Computador e Internet de las Cosas

Azar Miguel Augusto¹, Medrano José Federico², Tapia Maria Antonia³, Carlos Francisco Jorge⁴, Villafañe Juan Pablo⁵, Mamaní Arnaldo Ismael⁶, Aparicio Florentina⁷, Santos Walter Benjamín⁸, Estrada Bruno Rafael⁹ y Molloja Josué¹⁰

Facultad de Ingeniería - Universidad Nacional de Jujuy
Italo Palanca N° 10 (+54 0388 4221587) – San Salvador de Jujuy - Jujuy
auazar@live.com¹, jfedericomedrano@gmail.com², maria.antonía.tapia@gmail.com³,
fjcarlos1010@gmail.com⁴, juanpablovillaf@gmail.com⁵, arnaldomamani@yahoo.com.ar⁶,
fl17or@gmail.com⁷, waldo.unju@gmail.com⁸, bredostres@gmail.com⁹ y
josue.molloja@gmail.com¹⁰

Resumen

Actualmente dentro del campo de la informática existen diversas aplicaciones que se diseñan para funcionar en ambientes conectados al mundo exterior. Tal característica obliga a reenfocar el análisis y diseño de sistemas desde una perspectiva mas amplia que la convencional. Bajo este entorno el escenario se ha extendido hacia el diseño de aplicaciones sensibles al contexto (Context Aware). Esta amplitud exige que tanto los investigadores como los diseñadores incorporen nuevos elementos (sensores) para que efectivamente una aplicación pueda funcionar conectada tanto al medio ambiente como a las personas.

Si bien existen desde hace varias décadas aplicaciones que toman señales del exterior a través de sensores hacia sistemas embebidos, dichas aplicaciones no cuentan en su gran mayoría con algoritmos inteligentes (Artificial Intelligence) y menos aun con acceso a

Internet (Cloud Computing y Fog Computing).

El proyecto propone la investigación y el desarrollo de aplicaciones híbridas orientadas al uso de sensores, actuadores e interfaces mediante algoritmos y técnicas de inteligencia artificial. La implementación de este tipo de aplicaciones proporcionará mejoras en los sistemas sensibles al contexto actuales gracias al uso de técnicas basadas en IA y además constituirá un primer acercamiento hacia otros ámbitos tales como la robótica, el control de actuadores, brain computer interface, human computer interface, smartcities, smarhome, smartcar, smartgrid, artificial vision, augmented reality, internet of things, internet of mobile things, entre otros.

Palabras clave: Artificial Intelligence, Embeded Systems, Sensors, Brain-Computer Interface, Internet of Things.

Contexto

El proyecto de I/D/I presente se encuentra enmarcado dentro de la convocatoria realizada por la Secretaría de Ciencia y Técnica de Facultad de Ingeniería (UNJu) tramitada mediante Expediente S-7710/17 y refrendada mediante las Resoluciones C.A.F.I. N° 270/17 y C.A.F.I. N° 271/17. Dicha convocatoria estuvo orientada a docentes, incorporados recientemente al sistema de incentivos de la Secretaría de Políticas Universitarias, con categorías IV y V. A su vez, este llamado a la presentación de proyectos de investigación se encuentra financiado y da cumplimiento a los compromisos de mejoramiento establecido en la Resolución CONEAU N° 1230/12 para la acreditación de las Carreras de Ingeniería Informática y Licenciatura en Sistemas. El proyecto, luego del dictamen elevado por la Comisión Evaluadora, fue aprobado mediante Resolución C.A.F.I. N° 661/17.

En relación a los plazos establecidos, el proyecto inició el 01 de Enero de 2018 y finaliza el 31 de Diciembre de 2018.

Introducción

El uso de sensores como elemento adicional dentro del ámbito informático se ha incrementado en forma exponencial en los últimos años debido a su bajo costo y a la incipiente expansión de las tecnologías móviles. En este contexto solo en el mercado de sensores portátiles se espera un crecimiento del 30,14% hasta 2022 [1]. Del mismo modo, los sensores en la industria automotriz proyectan un incremento de 4 a 25 billones de dólares entre 2020 y 2030 [2]. Esta evolución es acompañada por una constante integración de tales sensores a los dispositivos móviles, el hogar, la

industria, las ciudades, el comercio, los automóviles, entre otros; estén o no estos conectados a Internet.

El uso de sensores también se ha extendido hacia el campo de Internet de las Cosas (IoT) en donde actualmente se concentran los esfuerzos, según Koreshoff [3], hacia tres categorías de investigación:

i) Explorar las ideas de diseño (servicios públicos, smarthome, salud /bienestar y datos personales/privacidad).

ii) Explorar sistemas a través del diseño (experimentos específicos dentro de un dominio, exploraciones relacionadas con actividades como el rastreo de personas, la automatización y la visibilidad de lo invisible, y exploraciones centradas en el sistema tales como las interfaces y sus componentes).

iii) Explorar componentes técnicos a través del diseño (consiste principalmente en: a) exploraciones centradas en el sistema, y b) exploraciones específicas del dominio).

Dentro del ámbito de IoT, si bien el uso de sensores cumple un rol fundamental, en [4] se amplifica el concepto señalando que son tres los componentes que se requieren para investigar dicho ámbito. Estos elementos básicos son: i) hardware integrado por sensores, actuadores, cámaras IP, CCTV y hardware de comunicaciones embebidas, ii) middleware: herramientas de almacenamiento e informática bajo demanda para el análisis de datos mediante Cloud Computing y el análisis de Big Data, iii) presentación: herramientas de visualización e interpretación fáciles de entender que se pueden diseñar para las diferentes aplicaciones.

Para Farooq y colegas [5], IoT utiliza una red de sensores (WSN, Wireless Sensor Network [6]) pero esa red forma

parte de una arquitectura de 6 capas: capa de código [7], capa de percepción [8], capa de red [9], capa middleware [10], capa de aplicación [11] y capa de negocios [12].

El análisis del estado del arte no solo implica centrarse en la importancia de los sensores dentro del ámbito de IoT, sino que no es necesario que tales sistemas se encuentren inmersos o conectados a la red Internet como es el caso de las Interfaces Cerebro-Computadora (BCI) [13] y las Interfaces Humano-Máquina (HCI) [14]. Entendiéndose como BCI a aquella interfaz constituida por un sistema de comunicación no muscular que no depende de las vías de salida normales del cerebro (nervios y músculos periféricos), por lo que puede proporcionar una conexión del cerebro a dispositivos de conmutación y control [15].

Por otra parte, la investigación en HCI actualmente involucra el desarrollo en el ámbito experiencial, entretenimiento y sus consecuencias psicológicas; siendo en menor medida su uso e incorporación en campo social [16].

Justificación del proyecto

Actualmente, si bien el desarrollo de rutinas algorítmicas orientadas al tratamiento de señales de sensores es ampliamente estudiado por el ámbito académico, existe una proporción exigua que analiza e investiga desde la óptica de la incorporación de algoritmos de inteligencia artificial para el tratamiento de los datos que dichos sensores proveen.

Además, las aplicaciones inteligentes permiten administrar eficientemente ciertos parámetros que otros algoritmos convencionales no lo consiguen con métodos matemáticos; como ser en el caso de algoritmos para la administración de la energía de la batería en un robot autónomo o en una sonda espacial debido

al corto tiempo en que sus celdas se encuentran expuestas a la energía solar.

Metodología

Los requerimientos de información partirán desde la base de las diferentes problemáticas existentes y planteadas por el conjunto de investigaciones en las que se describen las líneas de investigación que aun faltan por explorar. Se procederá al análisis correspondiente de las hipótesis de solución posibles y se recurrirá a una serie de experimentos en forma paralela que permitan un acercamiento a diversas soluciones.

Concretamente el enfoque que regirá como parte de la metodología de desarrollo es un paradigma específicamente diseñado para proyectos de IoT y se denomina: Ignite [17]. Dicho enfoque, propuesto por Slama y colegas en 2015, plantea (basado en una recopilación de las mejores prácticas de gestión y ejecución de proyectos IoT) dividir el desarrollo en dos fases: Iniciación de la estrategia y Entrega de soluciones. La primera, sesgada hacia la gestión comercial, se enfoca en definir la estrategia a seguir por parte de la empresa que desea migrar hacia proyectos de IoT, además de prepararla para adoptar dichos proyectos. Esto implica la creación y gestión de un portafolio de proyectos de IoT. La segunda, orientada al aspecto estrictamente tecnológico, direcciona a los gerentes de producto y jefes de proyecto en la planificación y ejecución. Se compone de tres fases las cuales son: Plan, Build y Run [18].

El esfuerzo está centrado en la segunda fase dado que especifica los pasos y detalles técnicos para el desarrollo de sistemas que involucran conectividad de sensores. Bajo este esquema el proyecto se divide en 2 partes: Iniciación del proyecto y Estructura del proyecto. En la primera se desarrollarán las siguientes

actividades: diseño de la solución inicial (análisis, proyección, planeamiento, diseño funcional y diseño técnico) y decisión entre construir y comprar (esto es, evaluar las soluciones comunes disponibles en el mercado, los recursos de hard y soft, el middleware IoT cloud/M2M y la comunicación wireless). Por otro lado, en cuanto a la estructura del proyecto se considerarán las siguientes actividades metodológicas: gestión del proyecto, tareas transversales, infraestructura solución y operaciones, servicios backend, servicios de comunicación, recursos de componentes y preparación de los recursos.

Líneas de Investigación, Desarrollo e Innovación

Los ejes que aborda el proyecto están en torno al uso de sistemas inteligentes para el tratamiento de las señales provistas por sensores a los sistemas embebidos.

Más específicamente, se centra en las siguientes temáticas: BCI, Artificial Vision, Robótica, IoT, AR y HCI.

Es por ello que se dispone de 3 computadoras industriales EDU-CIAA (pertenecientes a la Facultad de Ingeniería), computadoras de desarrollo (EDU-CIAA, Arduino, Raspberry PI) y sensores varios.

Resultados y Objetivos

El objetivo general es desarrollar algoritmos inteligentes sobre plataformas embebidas (online y standalone) para el tratamiento de señales provistas por sensores.

Objetivos particulares

- ✓ Analizar y evaluar las tecnologías existentes dentro del ámbito de los sensores y actuadores.
- ✓ Estudiar las alternativas de hardware/firmware/software para el tratamiento de los datos que los sensores proveen.
- ✓ Diseñar estructuras algorítmicas destinadas a mejorar las técnicas de inteligencia artificial convencionales.
- ✓ Seleccionar e incorporar una metodología de desarrollo adaptable a cada arquitectura y despliegue de hardware/software.

Con los conocimientos adquiridos se prevé dictar cursos-talleres en los que se transferirán las experiencias investigadas tanto a la comunidad educativa como al público interesado. Habrá un desarrollo y producción de conocimientos tanto en investigación básica como aplicada. En esta última se espera mayor impacto en la sociedad dado que se investigarán soluciones que mejoren su diario vivir a través del diseño de wearables, smartcities, smarthome, entre otros.

Formación de Recursos Humanos

El equipo de investigación está formado por un director, un co-director, dos egresados de Licenciatura en Sistemas, tres alumnos de Ingeniería Informática y tres alumnos de la carrera Analista Programador Universitario.

Se ha puesto énfasis en la formación de alumnos y es por eso que solo 2 docentes forman parte del equipo.

Referencias

- [1] Wearable Sensors Market by Type - Global Forecast to 2022. Online. <http://www.marketsandmarkets.com/Mar>

ket-Reports/wearable-sensor-market-158101489.html. Consultado el 10 de Octubre de 2017.

[2] Projected global market for autonomous driving sensor components. Online. <https://www.statista.com/statistics/423106/projected-global-market-for-autonomous-driving-sensor-components/>. Consultado el 13 de Octubre de 2017.

[3] Koreshoff T.L., Robertson T. y Leong T.W. Internet of Things: a review of literature and products. OzCHI '13 Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration. pp 335-344. Adelaide, Australia. 2013.

[4] Madakam S., Ramaswamy R. y Tripathi S. Internet of Things (IoT): A Literature Review. Journal of Computer and Communications. 3, 164-173. <http://dx.doi.org/10.4236/jcc.2015.35021>. 2015.

[5] Farooq M.U., Waseem M., Mazhar S., Khairi A. y Kamal T. A Review on Internet of Things (IoT). International Journal of Computer Applications (0975 8887). Volume 113, No. 1. 2015.

[6] Shen G. y Liu B. The visions, technologies, applications and security issues of Internet of Things. E-Business and E-Government (ICEE). pp. 1-4. 2011.

[7] Cheng X., Zhang M. y Sun F. Architecture of internet of things and its key technology integration based-on RFID. Fifth International Symposium on Computational Intelligence and Design. pp. 294-297. 2012.

[8] Bandyopadhyay D. y Sen J. Internet of Things - Applications and Challenges in Technology and Standardization. Wireless Personal Communications. Volume 58, Issue 1. pp. 49-69. 2011.

[9] Zhang Y. Technology Framework of the Internet of THings and Its

Application. Electrical and Control Engineering (ICECE). pp. 4109-4112. 2011.

[10] Shen G. y Liu B. The visions, technologies, applications and security issues of Internet of Things. E-Business and E-Government (ICEE). pp. 1-4. 2011.

[11] Wu M., Lu T., Ling F., Sun L. y Du H. Research on the architecture of Internet of things. Advanced Computer Theory and Engineering (ICACTE). pp. 484-487. 2010.

[12] Khan R., Khan S.U., Zaheer R. y Khan S. Future Internet: The Internet of Things Architecture, Possible Applications and Key Challenges. Proceedings of Frontiers of Information Technology (FIT). pp. 257-260. 2012.

[13] Rivas E.A., Rodriguez Morcillo C., Gianetti R. y Muñoz J.D. Dispositivos BCI. Tecnología actual y futuros desarrollos. <http://iies.es/dispositivos-bci-tecnologia-actual-y-futuros-desarrollos/>. 2016.

[14] Lazar J., Feng J.H. y Hochheiser H. Research Methods in Human-Computer Interaction. ISBN 978-0-170-72337-1. John Wiley and Sons Ltd publication. 2010.

[15] Choi I., Rhiu I., Lee Y., Yun M.H. y Nam C.S. A systematic review of hybrid brain-computer interfaces: Taxonomy and usability perspectives. PLoS ONE 12(4): e0176674. 2017.

[16] Hornbæk K. y Hertzum M. Technology Acceptance and User Experience: A Review of the Experiential Component in HCI. ACM Trans. Comput.-Hum. Interact. 24, 5, Article 33. 2017.

[17] Ignite | IoT Methodology. <http://enterprise-iot.org/book/enterprise-iot/part-ii-igniteiot-methodology/>. Consultado el 07 de Octubre de 2017.

[18] Slama D., Puhlmann F., Morrish J. y Bhatnagar R.M. Enterprise IoT, Strategies & Best Practices for Connected Products & Services. O'Reilly. 2015.

Monitoreo de carga por métodos no invasivos en el hogar argentino utilizando redes neuronales

Diego Cocconi*, Raúl Beinotti*, Rebeca Yuan*, Micaela Mulassano*, Javier Bruno*, Matías Beltramone*.

Departamento de Ingeniería en Sistemas de Información / Facultad Regional San Francisco /
Universidad Tecnológica Nacional (UTN)
Av. de la Universidad 501, 2400, San Francisco, Córdoba, Argentina, (03564) 431019 / 435402
*{dcocconi, rbeinotti, ryuan, mmulassano, jbruno, mbeltramone}@sanfrancisco.utn.edu.ar

Resumen

En la actualidad, mucha gente podría verse interesada en el monitoreo energético de sus viviendas, con el fin de optimizar sus consumos. De esta manera, se conocerían los artefactos eléctricos que más energía consumen, su incidencia en el tiempo, cuán representativo resulta el consumo del resto de los artefactos y aquellos artefactos que podrían estar fallando.

Utilizando Non-Intrusive Load Monitoring (NILM) y redes neuronales (del inglés Artificial Neural Networks, ANN), el presente proyecto propone ofrecer esta posibilidad.

Palabras clave: identificación de carga, monitoreo de energía, NILM, redes neuronales.

Contexto

El presente trabajo forma parte del proyecto de investigación I+D UTN 4881 “Monitoreo domiciliario utilizando redes neuronales a partir de una medición de energía totalizada (NILM)”. El mismo está homologado como proyecto de investigación y desarrollo en la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional. En el marco de dicho proyecto se propone implementar NILM mediante *machine learning* para el análisis del consumo de energía en los hogares argentinos, a través de la utilización de redes neuronales artificiales.

1. Introducción

En la actualidad, mucha gente podría verse interesada en el monitoreo energético de sus viviendas, con el fin de optimizar sus consumos. Dos enfoques fueron planteados para realizar tal monitoreo [1]: (1) la utilización de medidores independientes para cada artefacto eléctrico; (2) la aplicación de *Non-Intrusive Load Monitoring* (NILM) [2] o *Non-Intrusive (Appliance) Load Monitoring* (NIALM o NALM) [3]. NILM es una técnica computacional que a partir de una medida total de consumo de energía logra identificar los artefactos eléctricos individuales que se encuentran consumiendo la misma [4]; aunque midiendo cada artefacto puede resultar un método más exacto que NILM, las desventajas prácticas como elevados costos, múltiples configuraciones de sensores y complejidad en la instalación, favorecen el uso de esta técnica [5].

Un enfoque comúnmente utilizado para implementar NILM involucra las siguientes etapas: (1) *adquisición de datos*; (2) *extracción de features*; e (3) *inferencia o aprendizaje*. Durante esta última etapa, las diversas *features* de los artefactos eléctricos extraídas a partir de los datos de consumo son procesadas según diferentes algoritmos en orden de identificar los artefactos. Se suelen emplear técnicas supervisadas de *machine learning* en esta etapa, las cuales requieren datos etiquetados [5]; estas técnicas generalmente implican un proceso de aprendizaje lento y son vulnerables a cambios en el inventario de artefactos [1].

Por lo detallado anteriormente, se requieren abordajes que permitan obtener resultados más exactos, sin insumir tiempos prolongados de entrenamiento ni muchos recursos, e independientes de grandes inventarios de *features* de artefactos eléctricos.

2. Líneas de Investigación, Desarrollo e Innovación

Siguiendo la línea de investigación correspondiente a lo expuesto en este trabajo, se llevarán a cabo actividades relacionadas con las siguientes áreas temáticas:

- Ingeniería de Software.
- Mediciones Eléctricas.
- Eficiencia y Calidad de Energía Eléctrica.

3. Objetivos y Resultados Esperados

Resulta difícil para los usuarios de artefactos eléctricos determinar su consumo real de acuerdo a las condiciones de uso; en consecuencia, también es complejo determinar la incidencia de su uso a través del tiempo (diariamente, mensualmente, estacionalmente, etc.). Poder distinguir entre artefactos eléctricos con consumos importantes y aquellos que forman parte de un "piso" que no genera grandes consumos (como por ejemplo, *notebooks*, impresoras, etc.) puede resultar interesante para el usuario, básicamente por tres razones: (1) quedarían identificados los *artefactos que más consumen y su incidencia en el tiempo*, lo cual podría ayudar a moderar su uso; (2) se sabría si el *consumo de este "piso" puede resultar significativo*; y (3) podrían *detectarse aquellos artefactos con posibles fallas*.

Se plantea entonces ofrecer esta posibilidad al usuario de un modo que no implique dificultades ni complejidad adicional. Utilizando NILM la complejidad para el usuario se reduciría considerablemente; en tal caso, sería necesario identificar cada artefacto en una

medida de consumo total y determinar su consumo individual. Técnicas de *machine learning* como las *redes neuronales* (del inglés *Artificial Neural Networks*, ANN) han demostrado ser útiles para esta tarea. Sin embargo, la introducción de ANN conlleva desafíos adicionales: evitar que la arquitectura de la red resulte difícilmente entrenable, poco generalizadora ante nuevos artefactos y poco generalizadora para todos los comportamientos de un mismo artefacto. Entonces, se cuestiona que, ampliando los datos de la medición de consumo total con información proveniente de la extracción de *features* y utilizando dicho conjunto como entrada a la ANN, esto puede enriquecer la calidad de respuesta de la red, evitando problemas específicos inherentes a la misma.

En resumen, el principal objetivo de la línea de investigación propuesta es la identificación eficiente de los artefactos encendidos en la línea eléctrica de una vivienda mediante la aplicación de una ANN, pudiendo a través del tiempo determinar su consumo individual del modo más exacto posible, sin intervenir en el interior del hogar.

Para la consecución de este objetivo general, deberán tenerse en cuenta los siguientes objetivos específicos:

1. Modelar un sistema que permita la identificación de los artefactos eléctricos en uso dentro de una vivienda (los más importantes o representativos) a partir de la medida del consumo total de energía, utilizando una ANN.
2. Modelar un sistema que permita determinar el consumo de los artefactos identificados.
3. Modelar un sistema que permita determinar el consumo total de aquellos artefactos no tan importantes, no representativos o no identificados, que forman parte de un "piso".
4. Definir un sistema de almacenaje, recuperación y procesamiento de datos para proporcionar el consumo histórico de un artefacto y detectar fallas.
5. Definir un sistema de almacenaje, recuperación y procesamiento de datos

para determinar patrones de consumo (por ejemplo, patrones de consumo en invierno, en verano, etc.) de un artefacto.

- Proponer una arquitectura de ANN que permita satisfacer cuestiones referidas a la calidad de su respuesta: evitar que resulte difícilmente entrenable, poco generalizadora ante nuevos artefactos y poco generalizadora para todos los comportamientos de un mismo artefacto.

Para materializar el objetivo principal de este proyecto se propone una metodología “*design science research*” [13], pues provee un marco de trabajo ampliamente aceptado para realizar investigaciones relacionadas con sistemas de información, centrándose en resolver problemas por medio del diseño y desarrollo de artefactos.

La investigación comienza con un enfoque centrado en problemas, por lo que la etapa inicial consistió en definir problemas de investigación, hipótesis y objetivos. La segunda etapa implicará definir los artefactos para satisfacer los objetivos específicos.

Considerando el primer objetivo, se tendrán en cuenta tres fases: (1) *construcción del sistema NILM*; (2) *recolección y preparación de datos* para el entrenamiento de la ANN; y (3) *desarrollo de un entorno de pruebas* para comparar el desempeño de la red con otras alternativas conocidas.

La fase de construcción del sistema NILM precisará la elección de sensores de corriente y de voltaje; se analizarán *sensores de efecto Hall* y *sensores de efecto capacitivo*. Luego se evaluarán acondicionadores de señal apropiados (*amplificadores operacionales, amplificadores de instrumentación, etc.*). Seguidamente, se considerarán *convertidores AD* y entornos de desarrollo de sistemas electrónicos (*Arduino, Raspberry Pi, etc.*), conjuntamente con sus sistemas operativos soportados, que permitan implementar una microcomputadora. Se evaluarán luego lenguajes de programación (*C, Java, Matlab, Python*) y sus respectivos *frameworks/librerías* para implementar una función que permita la extracción de *features* y para implementar la ANN. Para mostrar la

información al usuario se tendrán en cuenta *pantallas LCD* y *conexión TCP/IP*. Un esquema de la arquitectura tentativa se ilustra en Fig. 1.

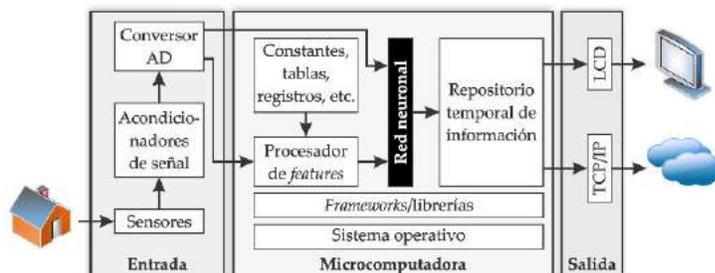


Fig. 1: Arquitectura del sistema NILM propuesto.

La segunda fase (recolección y preparación de datos) comenzará con la selección de un *analizador de potencia* o una *placa de adquisición de datos* que permita muestrear el consumo de un artefacto eléctrico en particular, con el nivel de detalle adecuado (baja frecuencia o alta frecuencia) y almacene dichos datos, para su posterior transmisión a una PC. El formato de almacenamiento de los datos en la PC será el utilizado por el MIT en el *dataset* de referencia REED¹. Una vez recopilados los datos para diversos artefactos, se procederá a su transformación en un esquema más adecuado para el entrenamiento de la red; se evaluarán bases de datos relacionales (*MySQL, SQL Server, PostgreSQL*). Una vez disponibles los datos en la base de datos, se considerarán diferentes lenguajes de programación (*C#, Java, Matlab, Python*) para identificar los ciclos de activación de cada artefacto, etiquetar los datos y definir una plataforma de entrenamiento para la ANN.

Las actividades de la tercera fase (desarrollo de un entorno de pruebas) incluyen utilizar la base de datos elegida y evaluar diferentes lenguajes de programación (*C#, Java, Matlab, Python*) para definir una plataforma de pruebas para la ANN, con el fin de comparar la alternativa propuesta con otros algoritmos conocidos y medir el desempeño utilizando diferentes métricas (*F1, precision,*

¹ <http://redd.csail.mit.edu/>

recall, accuracy) [4, 6].

Acerca del segundo objetivo, se desarrollará en el lenguaje de implementación utilizado en la ANN del primer objetivo, el algoritmo adecuado para calcular el consumo de cada artefacto eléctrico identificado.

Para el tercer objetivo, se desarrollará en el lenguaje de implementación utilizado en la ANN del primer objetivo, el algoritmo adecuado para calcular el consumo de aquellos artefactos eléctricos no tan importantes, no representativos o no identificados (“piso”).

Sobre el cuarto objetivo, se evaluarán bases de datos relacionales (*MySQL, SQL Server, PostgreSQL*), diferentes lenguajes de programación para desarrollo web (*HTML, JavaScript, PHP, Python*) y sus respectivos *frameworks* (*Laravel, Django*) para implementar una aplicación Web que permita comparar consumos históricos y determinar fallas en los artefactos eléctricos. En Fig. 2 se puede apreciar un esquema de la arquitectura tentativa del sistema Web.

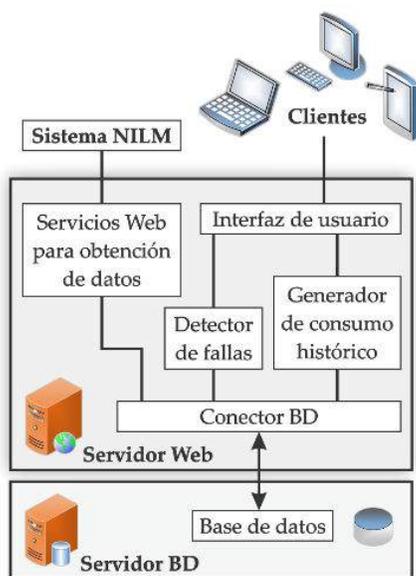


Fig. 2: Arquitectura del sistema Web propuesto.

Teniendo en cuenta el quinto objetivo, mediante la base de datos, el lenguaje de programación Web y su *framework* del cuarto objetivo, implementar una aplicación Web que permita determinar patrones de consumo de los artefactos.

Finalmente, para hacer frente al sexto objetivo, se evaluarán diferentes arquitecturas de ANN, sobre todo *deep neural networks: Restricted Boltzmann Machines (RBM), Deep Belief Networks (DBN), AutoEncoders (AE), deep Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN)* [4, 7]; además, se evaluarán diferentes *features* que podrían ayudar a mejorar el desempeño de la red.

La etapa final del proyecto consistirá en validar si los artefactos obtenidos permiten solucionar satisfactoriamente los problemas planteados.

4. Formación de Recursos Humanos

El grupo de esta línea de investigación está conformado por docentes y alumnos de las carreras de Ingeniería en Sistemas de Información, Electrónica y Electromecánica.

De los docentes involucrados en el proyecto, uno de ellos se encuentra cursando un Doctorado de Ingeniería con mención en Sistemas de Información, en la UTN Facultad Regional Santa Fe. Dos de las docentes dictan la cátedra Inteligencia Artificial (Ingeniería en Sistemas de Información) en la UTN Facultad Regional San Francisco, donde se encuentran realizando sus tesis para las Maestrías en Ingeniería de Software y Calidad de Software.

Como iniciativa del grupo, se prevé la capacitación y formación de recursos humanos, por medio de las siguientes actividades:

- Participación en cursos de actualización y posgrado en el área de estudio (*machine learning* e inteligencia artificial).
- Transferencia de conocimiento y resultados a otras áreas de la facultad y a la industria local.
- Incorporar el conocimiento adquirido en las cátedras referentes a la temática planteada.
- Ofrecer charlas informativas del desarrollo e implementación del proyecto a distintos sectores de la industria y cátedras afines a la investigación.

- Convocar e introducir a los alumnos de las carreras de Ingeniería en Sistemas de Información, Electrónica y Electromecánica a la realización de actividades de investigación y desarrollo.

5. Referencias

- [1] Xu, W., Dong, M., Meira, P., y Freitas, W. (2014, Julio). "An event window based load monitoring technique for smart meters". En: PES General Meeting| Conference & Exposition, 2014 IEEE (pp. 1-1). IEEE.
- [2] Hart, G. W. (1992). "Nonintrusive appliance load monitoring". Proceedings of the IEEE, 80(12), 1870-1891.
- [3] Bernard, T., y Marx, M. (2016). "Unsupervised Learning Algorithm using multiple Electrical Low and High Frequency Features for the task of Load Dis-aggregation". En: Proceedings of the 3rd international workshop on NILM.
- [4] Kelly, J., y Knottenbelt, W. (2015, Noviembre). "Neural nilm: Deep neural networks applied to energy disaggregation". En: Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environ-ments (pp. 55-64). ACM.
- [5] Zoha, A., Gluhak, A., Imran, M. A., y Rajasegarar, S. (2012). "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey". Sensors, 12(12), 16838-16866.
- [6] Makonin, S., y Popowich, F. (2015). "Nonintrusive load monitoring (NILM) performance evaluation". Energy Efficiency, 8(4), 809-814.
- [7] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., y Alsaadi, F. E. (2017). "A survey of deep neural network architectures and their applications". Neurocomputing, 234, 11-26.

Deep Neural Network para Análisis Acústico

García Mario Alejandro ✉¹, Rosset Ana Lorena², Eduardo Destéfanis¹, Cerruti Santiago¹, Moyano Miguel¹

¹Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN FRC)

²Universidad Nacional de Córdoba (UNC)

mgarcia@frc.utn.edu.ar

RESUMEN

La valoración de la calidad vocal mediante el análisis audio-perceptual es parte de la rutina clínica de evaluación de pacientes con trastornos de la voz. La debilidad de este método reside en la subjetividad y en la necesidad de que sea realizada por oyentes experimentados. Este proyecto tiene como objetivo la realización de una clasificación automática de la calidad vocal, valuada en la escala GRBAS, a través de características extraídas del análisis acústico de la señal y técnicas de aprendizaje automático. Particularmente, en este trabajo se muestran los resultados del cálculo de *shimmer* con un modelo de *deep learning*.

Palabras clave: *machine learning*, *deep learning*, *acoustic analysis*

CONTEXTO

Este trabajo de investigación se desarrolla en el marco del proyecto “Análisis acústico de la voz con técnicas de aprendizaje automático” (UTN3947) de la Universidad Tecnológica Nacional, Facultad Regional Córdoba y cuenta con la colaboración del Departamento de Investigación Científica, Extensión y Capacitación “Raquel Maurette”, Escuela de Fonoaudiología, Facultad de Ciencias Médicas, Universidad Nacional de Córdoba.

1. INTRODUCCIÓN

Se intenta reconocer, de forma automática, características del análisis acústico de la voz que permitan clasificar muestras de audio. El estudio se enfoca en la medición de la calidad vocal según la escala GRBAS. La clasificación se realiza aplicando principalmente modelos de *deep learning*, un subgrupo de técnicas del campo de aprendizaje automático (*machine learning*). Las grabaciones de la voz, la clasificación de los ejemplos y la validación de los resultados se realizan por especialistas en análisis de la voz de la Escuela de Fonoaudiología de la Universidad Nacional de Córdoba. El análisis acústico se lleva a cabo en conjunto (especialistas vocales e integrantes de UTN) y el modelado y desarrollo de los clasificadores por los integrantes de UTN.

GRBAS: La escala GRBAS es un método de valoración perceptivo-auditivo de la voz. Surge de la necesidad de estandarizar la valoración subjetiva y de interrelacionar los aspectos auditivos y fisiológicos de la producción vocal. Está basada en estudios del año 1966 de la *Japan Society of Logopedics and Phoniatrics* [1] y posteriormente divulgada y descripta por Minoru Hirano en el año 1981 [2]. Consiste en la valoración de la

fuente glótica a través de 5 parámetros que forman el acrónimo GRBAS:

G: (*Grade*) Grado general de disfonía.

R: (*Roughness*) Rugosidad, irregularidad de la onda glótica.

B: (*Breathiness*) Soplosidad, sensación de escape de aire en la voz.

A: (*Asteny*) Astenia, pérdida de potencia.

S: (*Strain*) Tensión, sensación de hiperfunción vocal.

Puede valorarse de dos maneras: a través de 4 grados, desde el 0 al 3 o mediante un valor en un rango continuo de 0 a 100. En ambas el 0 es ausencia de disfonía y el 3 o 100 implican disfonía severa. La escala fue mundialmente adoptada y validada en numerosos países [3-6]. Actualmente se utiliza en la investigación y de manera rutinaria en los consultorios de los profesionales que hacen clínica vocal. Sirve como metodología simple y al alcance de la mano para valorar la evolución pre-post tratamiento. La debilidad de este método reside en la subjetividad de la valoración de la voz y en la necesidad de que sea realizada por oyentes experimentados en la escucha y la disociación de los parámetros [7,8].

Análisis acústico: Existen otras formas de analizar la voz de manera más objetiva a través del análisis acústico. Éste consiste en la digitalización de la señal vocal y su análisis mediante gráficos como el Espectrograma, el espectro FFT (*Fast Fourier Transform*) o LPC (*Linear Predictive Coding*) y medidas numéricas de perturbación de la señal, como *Jitter*, *Shimmer* y HNR (*Harmonics to Noise Ratio*).

Para lograr una integración de la valoración subjetiva (GRBAS u otras escalas) con el análisis acústico, se han realizado numerosos

trabajos de correlación [9,10], algunos relacionados a la voz normal y otros a diferentes patologías. Por ejemplo, el trabajo de Nuñez Batalla, F. et al [11] es un referente y establece una relación entre el parámetro de Astenia del GRBAS y el Espectrograma de banda angosta.

Aprendizaje automático: El aprendizaje automático o *machine learning* es un campo de las ciencias de la computación que abarca el estudio y la construcción de algoritmos capaces de aprender y hacer predicciones. Estas predicciones se pueden tomar como una clasificación de los datos de entrada a partir del reconocimiento de patrones existentes en los mismos

Estado del arte: La aplicación de técnicas de *deep learning* es el estado del arte en el análisis automático de audio, con la detección de los fonemas pronunciados y la identificación de la persona que habla como objetivos principales [12-18], pero también utilizadas en detección de emociones, edad, género, etc.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación que se presenta en este trabajo tiene como objetivo predecir con técnicas de aprendizaje automático la medida acústica *shimmer*, asociada a la calidad vocal.

Shimmer:

Shimmer es una medida acústica relativa a las perturbaciones de amplitud de una señal. Las variaciones de este tipo en la voz humana son perceptibles al oído y permiten caracterizar ciertas propiedades, tanto de la voz, como de las personas que la emiten [19]. El valor de *shimmer* está asociado a la calidad vocal, estado de ánimo, edad y género de las personas. Existen numerosos trabajos de

investigación que utilizan, entre otras, la medida *shimmer* con objetivos que van desde la detección de patologías [19-21] hasta la mejora de interfaces humano máquina a través de la estimación de la intensionalidad de una frase hablada [22]. Con respecto a las voces sintetizadas, en [23] se determina que cierto nivel se *shimmer* aumenta el grado de naturalidad.

Hay variantes en el cálculo de *shimmer*. La versión elegida para este trabajo es la de Klingholz y Martin [23], también conocida como *Relative Shimmer*.

Metodología:

Los datos de entrada para el modelo de predicción son espectrogramas calculados sobre archivos de audio sintetizado.

Se generan datos de audio sin armónicos. Al igual que en [19] la modulación en amplitud de la voz se aproxima con una onda senoidal. La expresión para generar cada señal de audio $y(t)$ es:

$$y(t) = \frac{1}{1+k} \sin(\alpha + 2\pi f_0 t) (1 + k \sin(\beta + 2\pi f_{mod} t))$$

donde t es tiempo [seg], f_0 es la frecuencia de vibración glótica [Hz], f_{mod} es la frecuencia de modulación [Hz], k es la constante de sensibilidad en amplitud del modulador, α y β son constantes para manejar la fase de la señal a modular y de la señal moduladora respectivamente.

Para la generación de los datos de entrenamiento y test, se toman valores aleatorios con distribución uniforme. f_0 toma valores en el intervalo [200; 1000]Hz, f_{mod} en [5; 10]Hz, k en [0; 0,4], α y β en [0; 2π].

El espectrograma se calcula sobre 2 segundos de audio generado con 44100 muestras/seg. Para el cálculo se utiliza una ventana tipo

Tukey(0.25) de ancho = 256, lo que determina una estructura de forma 129 x 393 (frecuencia/tiempo) que contiene la densidad espectral de la señal.

El modelo de predicción es una *deep neural network*. Esta red se diseña en un proceso que genera modelos de complejidad ascendente, aproximando en primer lugar el valor de f_0 , k y f_{mod} individualmente a partir de los datos espectrales, luego *shimmer* en función de f_0 , k y f_{mod} , y por último, *shimmer* en función de los datos espectrales.

Se utilizan juegos de datos de 3000 muestras para entrenamiento, 500 para test y 500 para validación.

3. RESULTADOS OBTENIDOS

Se obtuvo un modelo neuronal con una capa de convolución y tres capas densamente conectadas (figura 1) capaz de aproximar *shimmer* con datos espectrales como entrada. El error cuadrático medio obtenido sobre los datos de test fue $MSE = 5.8 \times 10^{-5}$ [25].

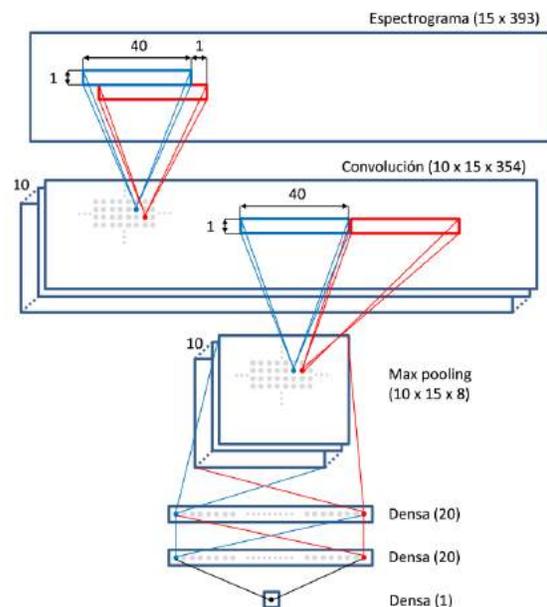


Figura 1. Modelo de *deep learning* para predicción de *shimmer*.

Se logró comprobar que para datos simples de audio modulados en amplitud por una onda senoidal, con parámetros de frecuencia fundamental, frecuencia moduladora y sensibilidad de modulación variables, es posible obtener un modelo neuronal capaz de aproximar el valor de *shimmer*. La importancia de este resultado radica en que, bajo las condiciones planteadas en el trabajo, se puede afirmar que un modelo de *deep learning* que respete la estructura del modelo presentado en sus primeras capas es capaz de utilizar el valor de *shimmer*, internamente calculado, para realizar clasificaciones de otro tipo, como la calidad vocal.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo del proyecto está formado por un docente/investigador de la UTN FRC, dos docentes/investigadores de la UNC y cuatro alumnos de la carrera de grado de la UTN FRC.

Además de formación de los alumnos participantes, el conocimiento generado por el proyecto se incorporará a las cátedras de los docentes de la UTN y UNC.

5. REFERENCIAS

[1] Isshiki, N., Yanagihara, N., & Morimoto, M. (1966). *Approach to the objective diagnosis of hoarseness*. *Folia Phoniatica et Logopaedica*, 18(6), 393-400.

[2] Hirano, M. (1981). *Clinical examination of voice* (Vol. 5). Springer.

[3] Yun, Y. S., Lee, E. K., Baek, C. H., & Son, Y. I. (2005). *The correlation of GRBAS scales and laryngeal stroboscopic findings for the assessment of voice therapy outcome in the patients with vocal nodules*. *Korean*

Journal of Otolaryngology-Head and Neck Surgery, 48(12), 1501-1505.

[4] Hui, H., Weijia, K., & Shusheng, G. (2007). *The Validation of Acoustic Analysis and Subjective Judgment Scales of Several Voice Disorders* [J]. *Journal of Audiology and Speech Pathology*, 3, 010.

[5] Karnell, M. P., Melton, S. D., Childes, J. M., Coleman, T. C., Dailey, S. A., & Hoffman, H. T. (2007). *Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders*. *Journal of Voice*, 21(5), 576-590.

[6] Jesus, L. M., Barney, A., Couto, P. S., Vilarinho, H., & Correia, A. (2009, December). *Voice quality evaluation using cape-v and GRBAS in european Portuguese*. In *MAVEBA* (pp. 61-64).

[7] Kreiman, J., & Gerratt, B. R. (2010). *Perceptual assessment of voice quality: past, present, and future*. *SIG 3 Perspectives on Voice and Voice Disorders*, 20(2), 62-67.

[8] Núñez-Batalla et al (2012). El espectrograma de banda estrecha como ayuda para el aprendizaje del método GRABS de análisis perceptual de la disfonía. *Acta Otorrinolaringológica Española*, 63(3), 173-179.

[9] Freitas, S. V., Pestana, P. M., Almeida, V., & Ferreira, A. (2015). *Integrating Voice Evaluation: Correlation Between Acoustic and Audio-Perceptual Measures*. *Journal of Voice*, 29(3), 390-e1.

[10] ELISEI, N. G. (2013). Percepción auditiva de voces patológicas. In XIV Reunión Nacional y III Encuentro Internacional de la Asociación Argentina de Ciencias del Comportamiento.

- [11] Nuñez Batalla, F., Corte Santos, P., Señaris Gonzalez, B., Rodriguez Prado, N., Suárez Nieto, C. (2004) Evaluación espectral de la hipofunción vocal. *Acta Otorrinolaringol. Esp.* 55:327-333.
- [12] Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., Kingsbury, B. (2012): Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *Signal Processing Magazine*, vol. 29.6, 82-97. IEEE.
- [13] Mitra, V., Sivaraman, G., Nam, H., Espy-Wilson, C., Saltzman, E., Tiede, M. (2017) Hybrid convolutional neural networks for articulatory and acoustic information based speech recognition. *Speech Communication*, vol. 89. pp 103-112.
- [14] Collobert, R., Puhersch, C., Synnaeve, G. (2016) Wav2letter: an end-to-end convnet-based speech recognition system. arXiv preprint arXiv:1609.03193.
- [15] Amodei, D., Ananthanarayanan, S., Anubhai, R., Bai, J., Battenberg, E., Case, C., Chen, J. (2016) Deep speech 2: End-to-end speech recognition in english and mandarin. *International Conference on Machine Learning*. pp. 173-182.
- [16] Palaz, D., Collobert, R. (2015) Analysis of cnn-based speech recognition system using raw speech as input (No. EPFL-REPORT-210039). Idiap.
- [17] Sainath, T. N., Kingsbury, B., Mohamed, A. R., Ramabhadran, B. (2013) Learning filter banks within a deep neural network framework. *IEEE Workshop onASRU*. pp 297-302. IEEE.
- [18] Farrús, M. (2007) Jitter and shimmer measurements for speaker recognition. 8th Annual Conference of ISCA. pp. 778-781. (2007)
- [19] Jafari, M., Till, J. A., Law-Till, C. B. (1993) Interactive effects of local smoothing window size and fundamental frequency on shimmer calculation. *Journal of Voice*, vol. 7.3. pp. 235-241.
- [20] Tsanas, A., Little, M. A., Fox, C., Ramig, L. O. (2014) Objective automatic assessment of rehabilitative speech treatment in Parkinson's disease. *Neural Systems and Rehabilitation Engineering*, IEEE Transactions, vol. 22.1. pp 181-190.
- [21] Gómez-Coello, A. et al. (2017) Voice Alterations in Patients With Spinocerebellar Ataxia Type 7 (SCA7): Clinical Genetic Correlations. *Journal of Voice*, vol. 31.1. pp. 123-e1.
- [22] Kotti, M., Patern, F. (2012) Speaker-independent emotion recognition exploiting a psychologically-inspired binary cascade classification schema. *International journal of speech technology*, vol. 15.2. pp. 131-150.
- [23] Yamasaki, R. et al (2017) Perturbation Measurements on the Degree of Naturalness of Synthesized Vowels. *Journal of Voice*, vol 31.3, 389-e1.
- [24] Klingholz, F., Martin, F. (1985) Quantitative spectral evaluation of shimmer and jitter. *J Speech Hear Res*, vol 28.2. pp 169-174.
- [25] García M.A., Destéfanis E.A. (2018) Deep Neural Networks for Shimmer Approximation in Synthesized Audio Signal. In: De Giusti A. (eds) *Computer Science – CACIC 2017*. CACIC 2017. Communications in Computer and Information Science, vol 790. Springer, Cham

Modelo basado en la Toma Decisiones con Criterios Múltiples para la elección de metodologías de Data Science

Karina B. Eckert¹, Paola V. Britos²

¹Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones

²Universidad Nacional de Río Negro

¹karinaeck@gmail.com, ²pbritos@unrn.edu.ar

RESUMEN

La capacidad de almacenamiento de datos generados por las organizaciones ha aumentado de manera significativa en las últimas décadas y poder analizarlos de manera adecuada, genera un factor estratégico para la Toma de Decisiones (TD). Se puede observar que en los últimos años se ha incrementado la cantidad de profesionales vinculados a la Ciencia de Datos (Data Science) y una de las habilidades requeridas es conocer y manejar a la perfección las metodologías disponibles y determinar cuál se adapta mejor para cada proyecto. Es por ello que al existir una diversidad de metodologías propuestas para el desarrollo de proyectos, en ocasiones la elección no es tarea sencilla, especialmente para los que se inician en el área. La presente investigación tiene como objetivo establecer un modelo basado en la Toma de Decisiones Multicriterios (o criterios múltiples), a través de métodos como el Proceso Analítico Jerárquico y su variante combinado con Lógica Difusa, con el propósito de establecer una base sólida para la selección de metodologías que guíen los proyectos de Minería de Datos o Data Science.

Palabras clave: Data Science, Metodologías de Minería de Datos, Toma de Decisiones Multicriterios, Proceso Analítico Jerárquico, Proceso Analítico Jerárquico Difuso.

CONTEXTO

El presente trabajo de investigación se desarrolla como trabajo final de posgrado de la Maestría en Tecnologías de Información, del tipo interinstitucional convenida entre la Facultad de Ciencias Exactas, Químicas y Naturales (FCEQyN) de la Universidad Nacional de Misiones (UNaM) y a la Facultad de Ciencias Exactas y Naturales y Agrimensura (FaCENA) de la Universidad Nacional del Nordeste (UNNE). El trabajo final de maestría se ha sido iniciado en la FCEQyN de la UNaM a mediados del año 2017 y se encuentra en etapa de desarrollo.

1. INTRODUCCIÓN

Con el avance de las Tecnologías de la Información y las Comunicaciones (TICs), los sistemas de almacenamiento poseen cada vez mayor capacidad de procesamiento, alto manejo y producción de datos y bajo costo del mismo; implicando que estos sistemas crezcan hasta límites inesperados. Es por ello, que es de trascendental que las organizaciones posean la capacidad de ser adaptativas y generar acciones a partir del conocimiento de la información almacenada.

En los últimos años han surgido una serie metodologías, procesos y herramientas, basadas en técnicas que facilitan el procesamiento avanzado de datos y permiten realizar un análisis en profundidad de los mismos de

manera automática; lo cual permite detectar la información oculta de los datos que a simple vista no son detectados y utilizarla de manera estratégica; esto es posible mediante la Ingeniería de Explotación de Información, popularmente conocida desde hace unas décadas como Minería de Datos o Data Mining (DM) y en la actualidad como Data Science (DS) [1].

Las técnicas de DM aportan información especializada destinada a favorecer una correcta TD en base a los resultados obtenidos a través de fuentes internas y concretas de las organizaciones, esto es, basarse en experiencias, datos, condiciones, entre otros factores. El resultado de este proceso es un conjunto de modelos o patrones, los cuales serán convertidos en información valiosa para la toma de decisiones. Los proyectos de DM pueden ser llevados a cabo en distintos escenarios, a partir de una situación organizacional o una exploración de los datos disponibles [1], [2], [3].

Al momento de seleccionar cuál de las metodologías disponibles actualmente utilizar para encarar un proyecto de DS, en ocasiones no es tarea sencilla, si bien existen algunos estudios comparativos sobre las mismas como los expuestos en [2], [3], [4], [5], [6], [7]; se propone como alternativa la utilización de la Teoría de la Toma de Decisiones Multicriterio (MCDM, Multiple Criteria Decision Making) [8], como soporte para la TD al momento de seleccionar una metodología de DS (o DM).

Los problemas de TD son procesos complejos en los cuales intervienen múltiples criterios; implican reconocimiento, análisis y evaluación de diversos aspectos; por lo cual es necesario utilizar herramientas que permitan discernir entre estos para obtener una solución que satisfaga en mejor grado la combinación de alternativas posibles [9]. Una de estas herramientas, es el Proceso Analítico Jerárquico (AHP, Analytic Hierarchy Process), creado por Saaty [10]. Los expertos hacen tres tipos generales de juicios para expresar su importancia, preferencia o probabilidad y

utilización, para elegir lo mejor entre las alternativas planteada; basan estos juicios en el conocimiento, en la memoria o en analizar los beneficios, costos y riesgos; sin normas se comparan las alternativas en lugar de calificarlas y las comparaciones deben estar dentro de un rango admisible de consistencia, cabe aclarar que AHP incluye tanto los métodos de calificación, como los de comparación. Por tanto, AHP permite estructurar, medir y sintetizar criterios para toma de decisiones con múltiple criterios [10], [11], [12]. El mismo ha sido aplicado ampliamente en la solución de una gran variedad de problemas, entre los cuales se pueden mencionar los presentados en [13], [14], [15], [16], [17], [18], [19], [20], [21].

Para seleccionar la mejor alternativa, AHP requiere que las comparaciones de los criterios y alternativas estén representadas en un número exacto [10], [11], usando una escala de nueve puntos, la cual representa las preferencias de los encargados de tomar las decisiones entre diferentes alternativas. Si bien la escala es simple y fácil de utilizar, no tiene en cuenta la incertidumbre asociada a los juicios humanos. Tal como afirmó Büyüközkan [22], "los responsables de la toma de decisiones generalmente se sienten mejor presentando sus juicios como un rango, en lugar de dar un valor exacto y fijo y lo anteriores porque él, ella o ellos son incapaces de explicar sus preferencias, dado el carácter difuso de comparación". Buckley [23] incorporó una matriz borrosa (difusa) al método AHP para representar la ambigüedad en las respuestas de las personas involucradas en la TD, lo cual proporciona un análisis para la TD con mayor validez. Huang y Wu [24] han expuesto que, "con la ayuda de la teoría de conjuntos difusos, se resuelven algunos defectos encontrados en el método tradicional de AHP, tales como la aplicación de escalas limitadas para la explicación de las consideraciones de expertos, la correlación entre los atributos para la TD, la imprecisión, la ambigüedad y la incertidumbre de expertos para encontrar los valores de las comparaciones" [24], [25]. Al igual que AHP tradicional,

Proceso Analítico Jerárquico Difuso (FAHP, Fuzzy Analytic Hierarchy Process) ha sido aplicado en diferentes escenarios, como los expuestos en [18], [26], [27], [28], [29].

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Los ejes de la investigación están relacionados a la aplicación de la teoría de MCDM, más precisamente AHP y FAHP, las cuales son ampliamente aplicadas para la selección de la mejor alternativa de las existentes, en función a la selección de múltiples criterios, para los cuales se utilizan etiquetas lingüísticas, que permiten evitar ambigüedades en las respuestas de los expertos. El modelo propuesto servirá como soporte en la toma de decisiones a la hora de seleccionar metodologías de Data Science (o DM), siendo este otro de los ejes de la investigación.

3. RESULTADOS Y OBJETIVOS

Objetivo General

Establecer un modelo basado en la MCDM, que sirva como soporte para la TD al momento de seleccionar una metodología de DS.

Objetivos Específicos

- Estudiar y analizar métodos de MCDM, como ser el AHP y FAHP.
- Revisar el estado del arte de las metodologías de DS (o DM) referentes en la actualidad.
- Definir criterios utilizando etiquetas lingüísticas para comparar de las metodologías de DS.
- Implementar el MCDM integrado a las etiquetas lingüísticas del caso de aplicación, evaluar y validar los resultados obtenidos con expertos del área.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación está inmersa en el desarrollo de un trabajo final de la Maestría en Tecnología de la Información dictada en la FCEQyN de la UNaM a cargo de la Ing. Karina B. Eckert y dirigida por la Dra. Paola V. Britos. Se prevé la incorporación de un alumno de grado como parte del equipo de trabajo, que pueda realizar su trabajo final de carrera dentro de la presente línea de investigación.

5. REFERENCIAS

- [1] K. B. Eckert and R. Suénaga, “Determinación del perfil académico a través del proceso de KDD”, ed. Editorial Académica Española, 2016.
- [2] J. M. Moine, “Metodologías para el descubrimiento de conocimiento en bases de datos: un estudio comparativo”, Tesis de Maestría, Facultad de Informática Universidad Nacional de La Plata, La Plata, Argentina, 2013.
- [3] M. T. Rodríguez Montequín, J. V. Álvarez Cabal, J. M. Mesa Fernández and A. González Valdés, “Metodologías para la realización de proyectos de Data Mining”, in VII Congreso Internacional de Ingeniería de Proyectos, pp. 257-265, 2005.
- [4] J. M. Moine, S. Gordillo and A. S. Haedo, “Análisis comparativo de metodologías para la gestión de proyectos de minería de datos”, in XVII Congreso Argentino de Ciencias de la Computación (CACIC), pp. 931-938, 2011.
- [5] A. Azevedo and M. F. Santos, “KDD, SEMMA and CRISP-DM: A parallel overview”, in IADIS European Conference on Data Mining, Amsterdam, 2008.
- [6] J. Giraldo Mejía and J., Jiménez Builes, “Caracterización del proceso de obtención de conocimiento y algunas metodologías para crear proyectos de Minería de Datos”, Revista Latinoamericana de Ingeniería de Software, vol. 1, no. 2, pp. 42-44, 2013.
- [7] P. Britos, “Procesos de explotación de información basados en sistemas inteligentes”, Tesis de Doctorado, Facultad

- de Informática Universidad Nacional de La Plata, La Plata, Argentina, 2008.
- [8] B. Roy Kolios, and R. Słowiński, “Questions guiding the choice of a multicriteria decision aiding method”, *EURO J Decis Process*, vol. 1, pp. 69-97, 2013.
- [9] M. Karanik, S. Gramajo, L. Wanderer, M. Giménez and D. Carpintero, “Multi-Criteria Decision Model based on AHP and Linguistic Information”, *Journal of Computer Science & Technology (JCS&T)*, vol. 14, no. 1, pp. 16-24, 2014.
- [10] T. L. Saaty, “How to make a decision: The analytic hierarchy process”, *European Journal of Operational Research*, vol. 48, no. 1, pp. 9-26, September 1990.
- [11] T. L. Saaty, “Decision making with the analytic hierarchy process”, *Int. J. Services Sciences*, vol. 1, no. 1, 2008.
- [12] V. Subramaniam and K. Lee, “Dynamic selection of dispatching rules for job shop scheduling”, *Production Planning and control*, vol. 11, no. 1, pp. 73-81, 2000.
- [13] J. C. Osorio Gómez and J. P. Orejuela Cabrera, “El proceso de análisis jerárquico (AHP) y la toma de decisiones multicriterio. Ejemplo de aplicación”, *Scientia et technica*, vol. 14, no. 39, 2008.
- [14] L. M. Ramírez, J. M. Flores and A. T. Vargas, “Análisis de Toma de Decisiones en Incubadoras Empresariales Mexicanas: un Modelo Basado en Jerarquización Analítica”, *CIAIQ 2017*, vol. 3, 2017.
- [15] M. A. Asma Bahurmoz, “The Analytic Hierarchy Process at Dar Al-Hekma, Saudi Arabia”, *Interfaces*, vol. 33, no. 4, pp. 70–78, 2003.
- [16] A. Başçetin, “An application of the analytic hierarchy process in equipment selection at Orhaneli open pit coal mine”, *Mining Technology*, vol. 113, no. 3, pp. 192 – 199, 2013.
- [17] D. Mendoza Casseres, “Aplicación de la teoría de decisión multicriterio discreta para ponderar factores en procesos de acreditación de alta calidad”, *Inge Cuc*, vol. 9, no. 1, pp. 25-41, 2013.
- [18] L. Vera Montenegro, “Aplicación y comparación de metodologías multicriterio (AHP y Fuzzi Logic) en la selección de tecnología postcosecha para pequeños productores de cacao”, Tesis de Doctorado, Universidad Politécnica de Valencia, Valencia, España, 2014.
- [19] E. Martínez Rodríguez, “Aplicación del proceso jerárquico de análisis en la selección de la localización de una PYME”, *Anuario Jurídico y Económico Escurialense*, pp. 523-542, 2007.
- [20] M. Lněnička, “AHP Model for the Big Data Analytics Platform Selection”, *Informatica Pragensia*, vol. 4, no. 2, pp. 108-121, 2015.
- [21] J. Wang, C. Li and S. Zhang, “Big Data Impact Analysis of Smart Grid based on AHP method”, *International Conference on Manufacturing Science and Engineering (ICMSE 2015)*, Atlantis Press, pp. 1487-1490, 2015.
- [22] G. Büyüközkan, C. Kabraman and D. Ruan, “A fuzzymulti-criteriadecision approach for software development strategy selection”, *Intern. Journal of General Systems*, vol. 33, no. 2, pp. 259-280, 2004.
- [23] J. J. Buckley, “Fuzzy Hierarchical Analysis”, *Fuzzy sets and systems*, vol.17, no.3, pp.233-247, 1985.
- [24] L. Huang and R. Wu, “Applying fuzzy analytic hierarchy process in the managerial talent assessment model -- an empirical study in Taiwan's semiconductor industry”, *International Journal of Technology Management*, pp. 105-130, 2005.
- [25] G. Herrera Enríquez, F. Guevara Viejó, S. Castillo Paez and D. Zambrano Vera, “Proceso Analítico Jerárquico Difuso en la selección de variables para la evaluación de la resiliencia en zonas afectadas por desastres”, *Ingeniería Industrial- Actualidad y Nuevas Tendencias*, vol. V, no. 16, pp. 45-66. 2016.

- [26] P. A. Barajas Díaz, V. A. Cloquell Ballester, J. Luisarcía Alcaraz and A. Alvarado Iniesta, “El Proceso Jerárquico Analítico y Lógica Difusa: Sus Aplicaciones”, in Congreso Internacional de investigación - Celaya, Gto. Academia Journals, vol. 4, no.3, 249-254, 2012.
- [27] C. A. Yajure, “Comparación de los métodos multicriterio AHP y AHP Difuso en la selección de la mejor tecnología para la producción de energía eléctrica a partir del carbón mineral”, *Scientia et Technica Año XX*, vol. 20, no. 3, pp.255-260, 2015.
- [28] R. Díaz, J. G. Piña, D. B. Ríos and M. P. Serafin, “Uso de AHP y Conjuntos Difusos para Mejorar la Toma de Decisiones. Caso: Selección de Empresas Contratistas de Construcción en la Administración Pública Venezolana”, in Seventh LACCEI Latin American and Caribbean Conference for Engineering and Technology (LACCEI’2009) - Energy and Technology for the Americas: Education, Innovation, Technology and Practice, 2009.

Diseño de Método de Ensamble Homogéneo para Clasificadores Débiles usando un esquema de reducción de datos simultaneo basado un enfoque co-evolutivo.

Ing. Corso Cynthia, Ing. Maldonado Calixto, Ing. Luque Claudio, Ing. Casatti Martín, Ing. Martínez Gimena.

Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información
Departamento Ingeniería en Sistemas de Información
Facultad Regional Córdoba/Universidad Tecnológica Nacional
Maestro M. López esq. Cruz Roja-Ciudad Universitaria-Córdoba
cynthia@bbs.frc.utn.edu.ar/calixto_maldonado@hotmail.com/cluque@prominente.com.ar/
m-casatti@gmail.com/gimemartinez05@gmail.com

RESUMEN

El objetivo de esta línea de investigación consiste en el diseño de una propuesta de método de ensamble, que permita mejorar la tasa de acierto para la resolución problemas de clasificación supervisada pertenecientes a distintos campos de aplicación. Más concretamente, esta propuesta pretende incorporar el desarrollo de una estrategia fundamentada en la búsqueda de atributos e instancias más significativos para el proceso de clasificación basado en enfoque evolutivo. El modelo resultante finalmente será aplicado para la clasificación de evento de fallos en equipos pertenecientes a un laboratorio de cómputos.

Palabras claves: *Métodos de ensamble, Bagging, Selección de instancias y atributos, Algoritmos evolutivos.*

CONTEXTO

Este trabajo pertenece al proyecto “Generación de Modelo Descriptivo para la caracterización de incidentes en equipos de un laboratorio de cómputos (Fase II)” PID-UTN3931.

Correspondiente al periodo de ejecución 2016-2018 del Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información (*GIDTSI*).

1. INTRODUCCIÓN

Con el propósito de mejorar la precisión en modelos de carácter predictivo, ha surgido un creciente interés en la definición de métodos que combinan hipótesis. Estos métodos se denominan multclasificadores, generan un conjunto de hipótesis e integran las predicciones del conjunto considerando un cierto criterio (normalmente por votación). La precisión obtenida por esta combinación, supera generalmente, la precisión de cada componente individual del conjunto [1].

La combinación de modelos se ha desarrollado principalmente para modelos predictivos, como la clasificación y regresión. Esta línea de investigación se focaliza en el estudio y análisis de modelos de multclasificación homogéneos para procesos de clasificación supervisada. En esta categoría existen diversos modelos que han sido propuestos, uno de ellos es Bagging.

Bagging (*Bootstrap Aggregating*) es un método de multclasificación para la optimización, en términos de precisión, de un modelo predictivo. Este método consiste en la creación de diferentes modelos de aprendizaje usando muestras aleatorias con reemplazo y luego combina los resultados obtenidos [2]. En la Figura 1 se visualiza el esquema de funcionamiento.

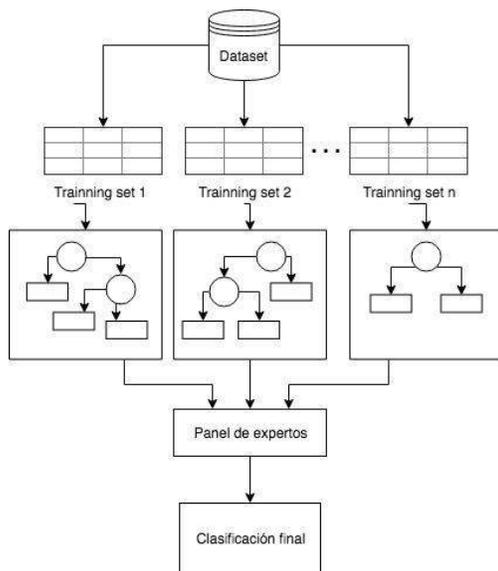


Figura 1. Esquema de algoritmo Bagging

Si bien el uso de métodos multclasificadores ha sido considerado una opción apropiada para mejorar la precisión de modelos de clasificación, es posible pensar que la integración de ciertas técnicas de preprocesamiento que permitan mejorar estos resultados.

En la práctica una dificultad que suele presentarse en el proceso de clasificación es el análisis de bases de datos de alta dimensionalidad. La causa de la alta dimensionalidad puede ocurrir por el aumento del número de instancias (conocidas como base de datos masivas) y de variables asociadas con cada

instancia (bases de datos con alta dimensionalidad). Una alternativa de solución a esta problemática es la posibilidad de conocer que atributos e instancias en la base de datos son realmente de utilidad para efectuar el proceso de clasificación.

La importancia del proceso de selección de características en cualquier problema de clasificación se pone de manifiesto puesto que permite eliminar las características que puedan inducir a error (características ruidosas), las características que no aporten mayor información (características irrelevantes) o aquellas que incluyen la misma información que otras (características redundantes) [3].

Aunque estos procesos de reducción de datos se definen por separado, como los mencionados anteriormente, es posible aplicarlos de manera simultánea. La Selección de Instancias y Atributos de manera simultánea o IFS (del inglés, *instance feature selection*) surge al no existir la definición de ningún criterio que permita decidir cuál método de reducción de datos ejecutar antes que otro.

En este sentido, diversas han sido las técnicas que abordan esta tarea desde el punto de vista de la computación evolutiva, considerando el uso de algoritmos genéticos aplicados a situaciones problemáticas de manera exitosa.

En [4] y [5] los autores presentaron un algoritmo genético para la realización de IFS, considerando su evaluación sobre un clasificador 1NN. Mientras que en [6] los autores presentaron el algoritmo IGA, que es un algoritmo genético inteligente que incorpora un operador de cruce ortogonal. Otros autores en [7] definen un algoritmo

genético híbrido (HGA) que reúne una serie de técnicas de búsqueda local y el propio algoritmo genético.

Uno de los trabajos más recientes presenta un modelo basado en algoritmos de co-evolución cooperativa que permite obtener tasas de error significativamente mejores que sus predecesores, al que denominaron IFS-CoCo [8]. IFS-CoCo consiste en una técnica wrapper [9], cuyo objetivo es maximizar la tasa de acierto del multclasificador y el porcentaje de reducción de instancias y atributos.

En este trabajo se define una adaptación del método de multclasificación Bagging considerando como algoritmo base a J48, basado en la integración de un enfoque de co-evolución cooperativa (IFS-CoCo), para el tratamiento de problemas de clasificación supervisada. En la Figura 2 se resume el esquema del método de ensamble propuesto.

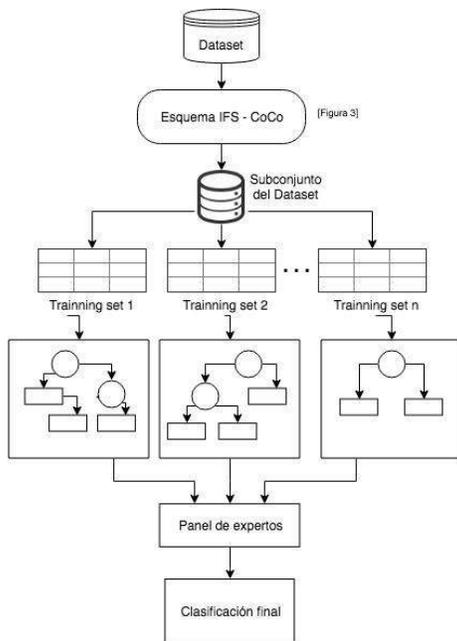


Figura 2. Esquema de método de ensamble propuesto.

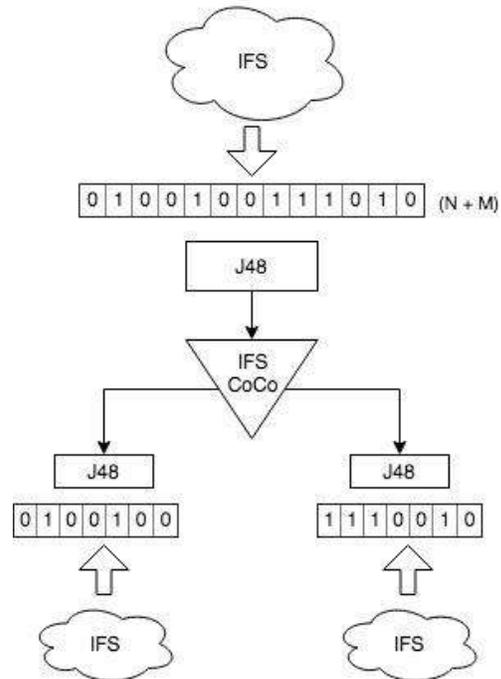


Figura 3. Esquema de reducción de IFS-Coco.

Este enfoque de co-evolución cooperativa se aplica en el proceso de selección de características e instancias que se implementan de manera simultánea en la base de datos inicial, con el propósito de obtener una configuración del conjunto de datos adecuada para efectuar el proceso de clasificación.

Finalmente, como objetivo a largo plazo se busca contribuir a la formación de recursos humanos en el ámbito del Aprendizaje Automático, la cual constituye un área científica de creciente interés.

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

La principal línea de investigación del proyecto es el área del Aprendizaje Automático, más precisamente la sub-área de los métodos de ensamble que

permiten la construcción de un conjunto de clasificadores cuyas decisiones son combinadas por un esquema específico, para la clasificación de nuevos ejemplos [10].

La arquitectura de estos métodos de ensamble puede ser homogénea o híbrida. En el primer caso se considera la utilización de un único algoritmo de minería de datos como base; mientras que en el caso de una arquitectura híbrida es posible la combinación de diferentes algoritmos como por ejemplo una red neuronal y una máquina de vector de soporte. Este trabajo de investigación se focaliza en el análisis y estudio de métodos de ensamble homogéneo considerando como base algoritmos de clasificación.

Existen técnicas de reducción de datos que son sumamente útiles, justamente para el caso de los algoritmos de clasificación que al presentar cambios estructurales en la base de datos inicial generan resultados muy diferentes en la clasificación final.

La innovación de este trabajo se ve reflejada en el diseño de un nuevo enfoque para un método de ensamble homogéneo que integra un esquema de reducción de atributos e instancias de forma simultánea sobre la base de datos inicial incorporando elementos de los algoritmos evolutivos.

3. RESULTADOS OBTENIDOS/ESPERADOS

Con esta línea de investigación se pretende principalmente lograr una contribución teórica referente a métodos de ensamble; mediante el diseño e

implementación de esta alternativa que combina de manera eficaz las ventajas propuestas de los algoritmos evolutivos para la selección simultánea de atributos e instancias más significativos en el proceso de clasificación.

En principio el proceso de selección de atributos e instancias de la base de datos original considerada en el modelo propuesto, se basa en una búsqueda bajo un enfoque coevolución cooperativa.

En líneas generales este enfoque de coevolución considera como punto de partida un conjunto de “ N ” instancias y “ M ” atributos. Cada cromosoma consiste en un número de genes, que es el representante de una característica o una instancia de la base de datos original. Esta propuesta considera tres poblaciones: i) población IS: cada gen representa una instancia. ii) población FS: cada gen representa una característica. iii) población IFS: los primeros “ N ” genes del cromosoma representan instancias, los genes restantes representan características (cromosoma de tamaño “ N ” x “ M ”). Cada una de ellas comparte la misma definición básica del cromosoma, que es una representación binaria. Al usar este esquema de representación, todos los cromosomas podrán definir un subconjunto de la base de datos inicial cuyo foco es la reducción de datos (características e instancias).

Con la integración de este esquema coevolutivo a un método de ensamble homogéneo como Bagging se espera obtener una mejora en términos de la tasa de acierto en la clasificación final.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto está conformado por docentes-investigadores pertenecientes a la carrera de grado de Ingeniería en Sistemas de Información. Todos los integrantes docentes del PID han participado del proceso de categorizaciones en investigación dentro del Programa de Incentivos del MECyT; así como en la categorización interna que posee la U.T.N.

Uno de los integrantes del proyecto está evaluando la posibilidad de iniciar su tesis de doctorado en la línea de investigación del citado proyecto.

Además participan alumnos avanzados en la carrera que realizan su práctica supervisada como requisito para el otorgamiento del título de grado de Ingeniero.

En este proyecto participan tres becarios, dos alumnos y un graduado que han logrado capacitarse mediante la ejecución de diversas tareas, complementando su formación académica con un acercamiento al ámbito de la investigación científica.

5. REFERENCIAS

- [1] J. Orallo Hernández, José Quintana, César Ramírez, “Introducción a la Minería de Datos”, pp.485-487, 2004.
- [2] Leo Breiman; “Bagging predictors, Machine Learning”, pp. 123–140, 1996.
- [3] Huan Liu and Hiroshi Motoda, “Computational Methods of Feature Selection” (Chapman & Hall/Crc Data Mining and Knowledge Discovery Series), Chapman & Hall/CRC, 2007.
- [4] Ludmila I. Kuncheva and Lakhmi C. Jain, “Nearest neighbor classifier: Simultaneous editing and feature selection,” Pattern Recognition Letters, vol. 20, no. 11-13, pp. 1149–1156, 1999.
- [5] T.Nakashima H.Ishibuchi and M.Nii, “Genetic algorithm- based instance and feature selection,” in Instance Selection and Construction for Data Mining, Motoda (Eds.), pp. 95–112, 2001.
- [6] Shinn-Ying Ho, Chia-Cheng Liu, and Soundy Liu, “Design of an optimal nearest neighbor classifier using an intelligent genetic algorithm,” Pattern Recogn; pp. 1495–1503, 2002.
- [7] Frederic Ros, Guillaou Serge, Marco Pintore, and Jacques R. Chretien, “Hybrid genetic algorithm for dual selection,” Pattern Anal. Appl., pp. 179–198, 2008.
- [8] Joaquín Derrac, Salvador García, and Francisco Herrera, “IFS-CoCo: Instance and feature selection based on cooperative coevolution with nearest neighbor rule,” Pattern Recogn., vol. 43, pp. 2082–2105, 2010.
- [9] L.J. Eshelman, “The CHC adaptive search algorithm: how to have safe search when engaging in nontraditional genetic recombination”, in: G.J.E. Rawlins (Ed.), Foundations of Genetic Algorithms; pp. 265–283, 1991.
- [10] María José Quintana Ramírez, José Hernández Orallo, “Extracción Automática de conocimiento en Base de Datos e Ingeniería de Software”, España, 2005.

Aprendizaje Automático. Aplicaciones en Visión por Computadora

L. Lanzarini¹, C. Estrebou¹, F. Ronchetti¹, F. Quiroga¹, C. Luna¹, R. Antonio¹, L. La Frazia¹, A. Rosete²

¹ Instituto de Investigación en Informática LIDI, Facultad de Informática, UNLP, La Plata, Argentina. *

² Universidad Tecnológica de La Habana “José Antonio Echeverría” (CUJAE), La Habana, Cuba

* Centro asociado de la Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CIC)

{laural, cesarest, fronchetti, fquiroga}@lidi.info.unlp.edu.ar
{carla.lunagennari, lucholafrazia}@gmail.com, ramiro.antonio@outlook.com,
rosete@ceis.cujae.edu.cu

CONTEXTO

Esta presentación corresponde al proyecto “Sistemas Inteligentes. Aplicaciones en Reconocimiento de Patrones, Minería de Datos y Big Data.” (Periodo 2018–2021) del Instituto de Investigación en Informática LIDI.

RESUMEN

Esta línea de investigación se centra en el estudio y desarrollo de Sistemas Inteligentes para la resolución de problemas de reconocimiento de patrones en imágenes y video, utilizando técnicas de Aprendizaje Automático. El trabajo presentado describe diferentes casos de aplicación en visión por computador de técnicas tanto supervisadas como no supervisadas.

Uno de los principales problemas desarrollados es el reconocimiento de lengua de señas. Este es un caso que presenta diversas aristas a atacar como el reconocimiento del intérprete, la segmentación de manos, la clasificación de diferentes configuraciones y de un gesto dinámico, entre otros problemas.

Con respecto a la segmentación de manos se realizaron diferentes trabajos, tanto utilizando marcadores de colores, redes neuronales capaces de reconocer el color de la piel de una persona, como así también redes convolucionales.

Por otro lado, para llevar a cabo la clasificación de diferentes gestos dinámicos, incluyendo la lengua de señas, se realizó un clasificador dinámico capaz de identificar

acciones humanas que faciliten la interfaz hombre/máquina.

En el área del procesamiento de video se está comenzando a investigar sobre detectores de peatones y automóviles para utilizar con cámaras instaladas en la vía pública.

Adicionalmente, se están realizando trabajos de clasificación de imágenes de especies de serpientes utilizando técnicas clásicas del aprendizaje automático

Palabras clave: Aprendizaje Automático, Visión por Computadoras, Redes Neuronales, Reconocimiento de Patrones, Lengua de Señas, Clasificación de Serpientes, Reconocimiento de Peatones y Autos.

1. INTRODUCCION

El Instituto de Investigación en Informática LIDI tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos métodos de Aprendizaje Automático. Los resultados obtenidos han sido medidos en la solución de problemas pertenecientes a distintas áreas.

En el III LIDI, desde hace varios años se viene trabajando en el procesamiento de señales de audio y video. Como resultado de estas investigaciones se han diseñado e implementado técnicas originales aplicables al reconocimiento tanto de gestos dinámicos como de detección de patrones en videos en diferentes problemas. En relación con esta línea, actualmente se están desarrollando los siguientes temas:

1.1. Reconocimiento de la lengua de señas

El reconocimiento de la lengua de señas es un campo de investigación relativamente nuevo cuyo objetivo final es traducir de la lengua de señas a una lengua escrita. Esto implica poder tomar un video en donde una persona habla en lengua de señas, y reconocer la posición de su cuerpo, su cara y sus manos, la expresión de su rostro, la forma de sus manos y también la de sus labios si la seña requiere pronunciar la palabra para desambiguar. Con esa información, se debe reconocer la seña realizada, para luego a partir de una secuencia de señas generar una traducción a una lengua escrita [6].

En esta área, se publicó un método probabilístico para clasificar señas en videos segmentados que abarca todas las etapas del reconocimiento [1, 2]. Este método probabilístico utiliza tres componentes esenciales en una seña: la posición de las manos, la forma y el movimiento. Cada componente es primero analizado por separado por sub-clasificadores para luego ser unificado en un clasificador global.

Este método no utiliza la información secuencial de la seña, es decir, no utiliza la información temporal. No obstante, los resultados de los experimentos muestran que aún con esa dificultad se pueden clasificar correctamente el 96% de las señas del conjunto de prueba.

Para llevar a cabo la validación de los métodos desarrollados, se utilizaron dos conjuntos de datos recolectados en el III-LIDI: LSA16 y LSA64. El primero, LSA16, contiene 800 imágenes con 16 clases de formas de mano y el segundo, LSA64, está formado por 3200 videos de 64 clases de señas dinámicas. Los detalles de la base de señas dinámicas LSA64 se encuentran publicados en [3].

Como resultado de esta línea de investigación se culminó exitosamente una tesis doctoral. Actualmente se están ampliando los métodos desarrollados para poder realizar una implementación de los métodos en un entorno real, lo que contribuiría con la comunidad para facilitar

la traducción entre la lengua de señas y el castellano.

1.2. Clasificación de formas de manos

Siguiendo con la temática de la sección anterior, una de las líneas investigadas es la clasificación de formas de manos. Las lenguas de señas utilizan un conjunto finito de formas de mano, que, en combinación con movimientos de las manos y el cuerpo, y expresiones faciales, se utilizan para señar [1].

En base a estudios previos [1,5], una etapa fundamental en el reconocimiento de la lengua de señas es la clasificación de estas formas de mano, y por ende un área prioritaria para mejorar el reconocimiento.

A su vez, las redes neuronales convolucionales, que han establecido nuevos estados del arte en casi todas las aplicaciones de visión por computadora, son idóneas para esta tarea. No obstante, no existe una revisión sistemática de la aplicabilidad de los diversos modelos de redes convolucionales a la clasificación de las formas de mano.

Por ende, se realizó una evaluación de desempeño de distintos modelos de redes convolucionales (LeNet, Inception, VGG, ResNets y AllConvolutional) con dos bases de datos de formas de mano (LSA16 [2] y RWTH-PHOENIX-Weather [5]).

Método	LSA16	RWTH
DeepHand (VGG)[5]	-	85.50
ProbSom[2]	92.30	-
Feedforward	86.58	60.27
LeNet	95.78	81.19
AllConvolutional	94.56	80.29
VGG	95.92	82.88
ResNet	93.49	80.89
Inception	91.98	75.33
Inception+SVM (pretrained)	93.67	78.12
Inception+NN (pretrained)	80.62	75.97

Tabla 1: Desempeño de cada método en LSA16 y RWTH-Phoenix-Weather. Los números representan porcentajes de ejemplos clasificados correctamente.

Tal como se muestra en la *Tabla 1*, se encontró que, salvo por Inception, todos los modelos consiguen un buen desempeño, incluso superando al estado del arte en LSA16. Además, se pudo observar que las redes convolucionales mejoran su desempeño al pre-segmentar la forma de la mano removiendo el fondo.

1.3. Localización de partes del cuerpo

Una etapa fundamental de todo el proceso de reconocimiento de un gesto en una imagen es la segmentación de las manos. En este sentido se están explorando distintas alternativas de segmentación de manos por el color de piel (sin marcadores) utilizando redes neuronales. En esta área se ha realizado un estudio del comportamiento de las redes neuronales RCE para segmentar la piel por color y su efectividad en diferentes sistemas de representación de color. Como resultado se llegó a la conclusión de que tanto el tiempo de cómputo como la efectividad de la segmentación son similares sin importar el sistema de representación de color elegido.

Este algoritmo de segmentación basado en la red RCE se ha utilizado en un prototipo de hardware y software que reconoce gestos dinámicos simples, en tiempo real, para controlar dispositivos electrónicos, en particular el control de un TV [7].

1.4. Detección de vehículos y peatones.

Continuando la línea de trabajos anteriores de reconocimiento de patentes, localización de objetos, y reconocimiento de acciones, se está desarrollando un proyecto en el que participan alumnos de grado de la Facultad de Informática para reconocer vehículos y peatones en cámaras de vigilancia.

El objetivo del proyecto es desarrollar la tecnología de base para un sistema de software que podría ser utilizado por los municipios para monitorear la entrada y salida de vehículos, controlar la velocidad de los mismos, y detectar posibles accidentes viales.

Con ese fin, se están evaluando distintos detectores de objetos, ya sean basados tanto en técnicas tradicionales como HOG y SVM, o en redes convolucionales como YOLO y R-CNN, utilizando bases de datos estándar como la Daimler Pedestrian Dataset o la Caltech Cars dataset.

1.5. Clasificación de especies de serpiente

En el marco de reconocimiento de objetos presentes en una imagen se está trabajando en conjunto con profesionales del CEPAVE¹ con el objetivo de desarrollar e implementar un algoritmo capaz de identificar la especie a la que pertenece una serpiente a partir de la imagen de un ejemplar. Interesan especialmente las especies de serpientes de la provincia de Buenos Aires y en particular las del partido de La Plata. Esto, si bien reduce considerablemente el número de opciones, no invalida la importancia y aplicabilidad de las investigaciones realizadas.

Cabe destacar la importancia de esta aplicación ya que permite determinar la peligrosidad de una serpiente. Por un lado, permite tomar decisiones rápidas ante un incidente y por otro lado contribuye a la conservación de las especies ya que las personas ante la duda de su peligrosidad deciden matarlas.

En esta aplicación en desarrollo, se han utilizado los algoritmos SIFT, SURF y ORB para extraer los descriptores locales de la imagen de una serpiente. Se han analizado distintas métricas de comparación de características para determinar la correspondencia entre los descriptores. Los resultados preliminares arrojan una tasa de acierto superior al 70% pero aún queda trabajo por realizar.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

- Estudio de técnicas de segmentación de objetos en movimiento presentes en un

¹ CEPAVE Centro de Estudios Parasitológicos y de Vectores - Conicet La Plata – U.N.L.P.

video.

- Representación y clasificación de configuraciones de manos para el lenguaje de señas.
- Clasificación de señas dinámicas.
- Estudio y análisis de las distintas representaciones de color.
- Redes neuronales competitivas dinámicas. Redes neuronales RCE.
- Detección y extracción de características. Puntos de interés y descriptores de regiones.
- Estudios de performance de los algoritmos desarrollados

3. RESULTADOS OBTENIDOS/ESPERADOS

- Desarrollo de un modelo de clasificación de señas segmentadas y comparación de su desempeño con otros modelos del estado del arte.
- Evaluación de desempeño de varias arquitecturas de redes convolucionales del estado del arte para la clasificación de formas de manos.
- Desarrollo de un método de clasificación de especies de serpientes basados en descriptores SIFT, SURF y ORB.
- Detección y clasificación de peatones y vehículos mediante cámaras RGB.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 2 profesores con dedicación exclusiva, 2 becarios de investigación UNLP con dedicación docente, 1 becario CIN, 2 tesis y 1 profesor extranjero.

Dentro de los temas involucrados en esta línea de investigación, en el último año se han finalizado 2 tesis de doctorado, y 2 tesinas de grado de Licenciatura.

Actualmente se están desarrollando 1 tesis de doctorado, 1 tesis de especialista y 3 tesinas de grado de Licenciatura. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

- [1] Ronchetti F., Quiroga F., Estrebow C., Lanzarini L., Rosete A. *Sign language recognition without frame-sequencing constraints: A proof of concept on the argentinian sign language*. Publicado en Ibero-American Conference on Artificial Intelligence IBERAMIA 2016 (pp. 338-349)
- [2] Ronchetti, F., Quiroga, F., Estrebow, C.A., Lanzarini. *Handshape recognition for Argentinian Sign Language using ProbSom*. Journal of Computer Science & Technology, vol. 16, N° 1, págs. 1-5, ISSN 1666-6038, 2016.
- [3] Ronchetti, F., Quiroga, F., Estrebow, C.A., Lanzarini, L.C., Rosete, A. . *LSA64: An Argentinian Sign Language Dataset*, publicado en el XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016) (pp. 794-803).
- [4] Quiroga, F., Antonio, R., Ronchetti, R., Lanzarini, L., Rosete, A. *A Study of Convolutional Architectures for Handshape Recognition applied to Sign Language*, publicado en el XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017) (pp. 13-22).
- [5] Koller, O., Ney, H., Bowden, R. . *Deep Hand: How to Train a CNN on 1 Million Hand Images When Your Data Is Continuous and Weakly Labelled*. Computer Vision and Pattern Recognition Conference (CVPR 2016) (pp. 3793-3802).
- [6] Cooper H., Holt B., and Bowden R. *Sign language recognition*. In Visual Analysis of Humans: Looking at People, chapter 27, Springer, 2011. (pp539- 562).
- [7] Luna Gennari C., Estrebow C., Lanzarini, L. *Reconocimiento de gestos aplicado al control de dispositivos*, publicado en el XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017) (pp.1040-1049).

Redes Inteligentes de Agua: Factores y métodos para la predicción del consumo residencial de agua potable

Pandolfi Daniel¹, Alba Enrique³, Andrea Villagra¹, Leguizamón Guillermo^{1 2}
dpandolfi@uaco.unpa.edu.ar, eat@uma.es, avillagra@uaco.unpa.edu.ar, legui@unsl.edu.ar

¹⁰Laboratorio de Tecnologías Emergentes (LabTEM) Instituto de Tecnología Aplicada (ITA)
Unidad Académica Caleta Olivia - Universidad Nacional de la Patagonia Austral

²Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC)
Departamento de Informática - Universidad Nacional de San Luis

³Grupo NEO, Departamento de Lenguajes y Ciencias de la Computación - Universidad de Málaga

Resumen

El concepto de ciudades inteligentes y sostenibles se introdujo para superar los grandes desafíos que enfrenta el desarrollo urbano. Por lo tanto, el crecimiento demográfico y el cambio climático presentan nuevos retos para las administraciones gubernamentales y las empresas de consumo de agua potable. Los sistemas inteligentes de distribución de agua potable deben ser cada vez más eficientes y responder a múltiples restricciones de abastecimiento. Las variables relacionadas a la demanda residencial de agua es uno de los parámetros más difíciles de determinar para modelar en este tipo de sistemas. El objetivo de esta línea es *identificar* los distintos factores que inciden en el estudio de la demanda de consumo de agua potable a escalas espaciales y temporales, e incorporar *modelos que proporcionen estimaciones confiables para predecir* el consumo de agua urbana a nivel residencial a corto, medio y largo plazo.

Palabras clave: Smart City, Big Data, Redes inteligentes de agua, Machine Learning.

Contexto

La línea de trabajo se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEM), Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Caleta Olivia Universidad Nacional de la Patagonia Austral, en el marco del Proyecto de Investigación 29/B255 “Soluciones inteligentes para el desarrollo urbano sostenible”. Este proyecto se desarrolla en cooperación con el LIDIC de la UNSL, y el Grupo NEO de la UMA (España).

Introducción

El agua desempeña un papel crucial en la ubicación y el crecimiento de las comunidades y el Foro Económico Mundial anunció en 2015 que la crisis del agua ocupa el octavo riesgo mundial con la mayor probabilidad de ocurrir dentro de los próximos 10 años [WEF15]. La demanda de agua urbana es parte de un sistema complejo, que depende de los patrones y procesos que surgen a través de interacciones de variables a múltiple escala.

House-Peters and Chang describen en [HPC11] una amplia revisión de la capacidad de estimar la demanda de agua en múltiples escenarios climáticos, de crecimiento poblacional y relacionándolos con el modelado de procesos hidrológicos urbanos. Para el análisis de demanda residencial de agua los autores [HPC11] han descrito distintos tipos de factores temporales y espaciales. Entre los factores temporales se pueden citar variables tales como temperatura, precipitaciones, velocidad del viento, evaporación, precio del agua, crecimiento de la población, crecimiento de la población y nivel de ingresos. En los factores espaciales se pueden mencionar variables tales como edad, tamaño de la familia, educación, tamaño de las viviendas, número de habitaciones, tamaño del espacio al aire libre, piscinas, proporción de viviendas unifamiliares, tipos de vivienda, índice de diferencia de vegetación normalizada, índice de calor urbano, y políticas ambientales.

Según Corbella y Sauri [CS09] los sistemas de distribución de agua atienden a la demanda de clientes residenciales, industriales, comerciales, institucionales y públicos. Las demandas se ven afectadas tanto por factores de impacto a largo plazo como cambio poblacional, condiciones económicas y políticas de conservación del agua y otros factores de impacto a corto plazo, incluidos los patrones climáticos estacionales y las demandas pico asociadas (verano).

El comportamiento de los consumidores puede facilitar la determinación de un enfoque más proactivo de la gestión de la demanda de agua y además el desarrollo de una estrategia de intervención que logren predicciones fiables en el consumo de agua doméstica.

Trabajos recientes describen el estado de las investigaciones, particularmente para los aspectos relacionados con los modelos predictivos describimos a continuación algunas de las comunicaciones más recientes. Walker et al. [WC+15] presentan un modelo de predicción de consumo doméstico de agua basado en redes neuronales, el modelo se basa en datos reales recopilados a partir de medidores inteligentes en tiempo real. En [SY+15] se presenta el resultado de la investigación de la encuesta como parte del Sistema de Apoyo Integrado para el Uso Eficiente del Agua y la Gestión de Recursos (ISS-EWATUS) financiado por la Unión Europea (UE). El trabajo incluye el análisis de tres elementos principales para identificar para el comportamiento de los consumidores de agua domésticos: comportamientos de uso final; características sociodemográficas y de propiedad; y construcciones psicosociales tales como actitudes y creencias. Chen et al. [CY+15] proponen un modelo de evaluación comparativa para el consumo doméstico de agua basado en *Adaptive Logic Networks* (ALN), los datos del mundo real son recopilados por un sistema de control del consumo de agua instalado en Sosnowiec (Polonia) y Skiathos (Grecia). En [CY+15] los autores presentan un trabajo basado en *Deep Learning* (DL) y redes neuronales artificiales (ANN) donde muestran las aplicaciones para la simulación, la optimización y el control de la operación de los sistemas de distribución de agua. En [CV+15] se presenta una metodología para evaluar la eficiencia del uso doméstico de agua basado en la comparación entre pares. Los grupos de pares se establecieron a través de análisis de conglomerados según sus características sociodemográficas más

relevantes. Sanz y Pérez [SP15] describen una metodología de colocación de sensores basada en el análisis de presión y sensibilidad al flujo utilizando el método *Singular Value Decomposition* (Descomposición del Valor Singular) y presentan una comparación de calibración de demanda en una red real con datos sintéticos. Shabani y Naser [SN15] utilizan una amplia gama de variables explicativas en sus modelos de previsión de demanda de agua. En [VC+15] se presenta una metodología novedosa para evaluar la eficiencia del uso de agua doméstica en interiores basada en funciones de evaluación del rendimiento que convierten variables de estado (por ejemplo, caudal, volumen) en índices de rendimiento que van desde 0 (el rendimiento más bajo) hasta 300 (excelente rendimiento). [SR+15] Salleh et al. presentan un estudio donde se investiga el efecto de las variaciones en los datos de consumo de agua en la clasificación de los niveles de uso doméstico de agua. En [duPJ15] du Plessis y Jacobs analizan el uso de agua al aire libre de propiedades residenciales donde las componentes de uso de agua definidas matemáticamente y combinadas para desarrollar un modelo predictivo. En [HK15] se presenta una investigación preliminar sobre la aplicación de un enfoque de pronóstico de demanda probabilístico para identificar roturas de tuberías. El método produce un pronóstico probabilístico de la demanda futura en condiciones normales. Candelieri et al. [CS+15] proponen un algoritmo de autoaprendizaje basado en *Support Vector Machine* completamente adaptable, basado en datos, para pronosticar la demanda de agua a corto plazo y con periodicidad horaria, basados en la disponibilidad de lectores automáticos de medición (AMR). En [CY+15] los autores presentan un trabajo basado en

Deep Learning (DL) y redes neuronales artificiales (ANN) donde muestran las aplicaciones para la simulación, la optimización y el control de la operación de los sistemas de distribución de agua.

Líneas de investigación y desarrollo

En esta sección se describe la línea de investigación que se lleva a cabo en el proyecto y la línea de trabajo con medio ambiente y redes inteligentes de agua.

Una ciudad inteligente es la que realiza actividades en al menos una de las seis áreas de acción inteligente: Economía inteligente, Sociedad inteligente, Gobernanza inteligente, Movilidad inteligente, Medioambiente inteligente, Modo de vida inteligente. Algunas aplicaciones que se pueden definir en estas áreas y que son de interés en este trabajo de investigación son: (a) Movilidad: rutas dinámicas personalizadas; seguridad urbana y (b) Medioambiente: gestión inteligente de la basura; agua inteligente.

La gestión de redes inteligentes de agua involucran muchas decisiones a nivel operacional tomando como base factores de decisión tales como el control de válvulas y bombas, la predicción de consumo, análisis de las variables de demanda y consumo residencial. Particularmente, esta línea de trabajo se enfocará a la predicción de la demanda de agua para el consumo residencial a tres escalas de tiempo *nowcasting*, *forecasting*, *foresight*.

Resultados obtenidos/esperados

Como objetivo general se espera modelar, resolver e implementar un amplio y variado conjunto de servicios inteligentes de la ciudad y producir un mayor impacto en Ciencia e Industria.

Como objetivos específicos se pretende: (a) Estudiar el estado de tecnologías y técnicas de Machine Learning, para resolver problemas en problemas de Smart-Water. (b) Proponer y seleccionar tecnologías a los problemas identificados. (c) Diseñar y construir prototipos de solución a los problemas identificados.

Formación de recursos humanos

Esta línea de investigación proporcionará un marco propicio para la iniciación y/o finalización de estudios de posgrado de los integrantes docentes. De igual forma, será un ámbito adecuado para la realización de tesis de grado. En ese sentido, dos integrantes de este proyecto de investigación está desarrollando su Tesis de Maestría en temáticas afines y un integrante está desarrollando su Tesis de doctorado. Además, se cuenta con un becario alumno de la carrera de Ingeniería en Sistemas.

Referencias

[CS+15] Candelieri A., Soldi D., Archetti F., Short-term forecasting of hourly water consumption by using automatic metering readers data, *Procedia Engineering*, Volume 119, 2015, Pages 844-853.

[CV+15] Catarina J., Vieira P., Rebelo M., Covas D., Assessment of Water use Efficiency in the Household Using Cluster Analysis, *Procedia Engineering*, Volume 119, 2015, Pages 820-827.

[CY+15] Chen X., Yang S.H., Yang L., Chen X., A Benchmarking Model for Household Water Consumption Based on Adaptive Logic Networks, *Procedia Engineering*, Volume 119, 2015, Pages 1391-1398.

[CS09] Corbella, H. M., and Sauri Pujol D. (2009), What lies behind domestic water use? A review essay on the drivers of domestic water consumption, *Bol. Asoc. Geogr. Esp.*, 50, 297-314.

[duPJ15] du Plessis J.L., Jacobs H. E., Procedure to Derive Parameters for Stochastic

Modelling of Outdoor Water use in Residential Estates, *Procedia Engineering*, Volume 119, 2015, Pages 803-812.

[HPC11] House-Peters, L. A., and Chang H. (2011), Urban water demand modeling: Review of concepts, methods, and organizing principles, *Water Resources Research*, vol. 47, W05401.

[HK15] Hutton C., Kapelan Z., Real-time Burst Detection in Water Distribution Systems Using a Bayesian Demand Forecasting Methodology, *Procedia Engineering*, Volume 119, 2015, Pages 13-1.

[SP15] Sanz G., Pérez R., Comparison of demand calibration in water distribution networks using pressure and flow sensors, *Procedia Engineering*, Volume 119, 2015, Pages 771-780, ISSN 1877-7058.

[SR+15] Salleh N., Rasmani K. A., Jamil N. I., The Effect of Variations in Micro-components of Domestic Water Consumption Data on the Classification of Excessive Water Usage, *Procedia - Social and Behavioral Sciences*, Volume 195, 2015, Pages 1865-1871.

[SN15] Shabani S., Naser Gh., Dynamic Nature of Explanatory Variables in Water Demand Forecasting, *Procedia Engineering*, Volume 119, 2015, Pages 781-787.

[SY+15] Shan Y., Yang L., Perren K., Zhang Y., Household Water Consumption: Insight from a Survey in Greece and Poland, *Procedia Engineering*, Volume 119, 2015, Pages 1409-1418.

[VC+15] Vieira P., Catarina J., Covas D., Novel Performance Assessment Indices for Domestic Water Use, *Procedia Engineering*, Volume 119, 2015, Pages 813-819.

[WC+15] Walker D., Creaco E., Vamvakieridou-Lyroudia L., Farmani R., Kapelan Z., Savić D., Forecasting Domestic Water Consumption from Smart Meter Readings Using Statistical Methods and Artificial Neural Networks, *Procedia Engineering*, Vol 119, 2015, Pages 1419-1428.

[WEF15] World Economic Forum (2015). *Global Risks 2015 Report*. Jan, 2015.

[WE+15] Wu Z. Y., El-Maghraby M., Pathak S., Applications of Deep Learning for Smart Water Networks, *Procedia Engineering*, Volume 119, 2015, Pages 479-485.

Soluciones inteligentes para el desarrollo urbano sostenible

Villagra A.¹, Errecalde M.^{1,2}, Molina D.¹, Varas V.¹, Orozco S.¹, Valdéz J.¹, Rasjido J.¹, Mercado V.¹, Montenegro C.¹, Carballo L.¹, Pandolfi D.¹

¹Laboratorio de Tecnologías Emergentes (LabTEM)
Instituto de Tecnología Aplicada (ITA) - Unidad Académica Caleta Olivia
Universidad Nacional de la Patagonia Austral

²Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC)
Departamento de Informática - Universidad Nacional de San Luis

avillagra@uaco.unpa.edu.ar, merreca@unsl.edu.ar, {dmolina, vvaras, sorozco, jcvldez, jrasjido, vmercado, cmontenegro, [icarballo.dpandolfi](mailto:icarballo.dpandolfi@uaco.unpa.edu.ar)}@uaco.unpa.edu.ar

RESUMEN

Una ciudad inteligente es una ciudad innovadora que utiliza Tecnologías de la Información y Comunicación (TIC) y otros medios para mejorar la toma de decisiones, así como la eficiencia de las operaciones, los servicios urbanos y su competitividad, a la vez que se garantiza la atención a las necesidades de las generaciones actuales y futuras en relación con los aspectos económicos, sociales y medioambientales. Una ciudad Inteligente integra sus diferentes áreas utilizando redes de comunicación de banda ancha, computación en nube, dispositivos inteligentes móviles, programas de análisis y sensores, permitiendo construir y aplicar el conocimiento para apoyar la toma de decisiones y ofrecer mayor calidad de vida y beneficios a sus ciudadanos.

En esta línea la hipótesis de trabajo es que se puede resolver varios problemas diferentes analizando sus características cuantitativas y cualitativas subyacentes,

así como proporcionando algoritmos avanzados que pueden buscar, optimizar y aprender por sí mismos en aquellas situaciones donde el conocimiento del problema es muy limitado. En particular, en este proyecto se propone identificar, resolver y realizar prototipos de aplicaciones y servicios en dos ejes movilidad y medio ambiente.

Palabras clave: Ciudades Inteligentes, Metaheurísticas, Big Data, Redes de sensores inalámbricos.

CONTEXTO

La línea de trabajo se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEM), Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Caleta Olivia Universidad Nacional de la Patagonia Austral, en el marco del Proyecto de Investigación 29/B255 “Soluciones inteligentes para el desarrollo urbano sostenible”.

1. INTRODUCCIÓN

Vivimos en la convergencia de dos fenómenos importantes en la historia de la humanidad: la aceleración de la urbanización a nivel mundial y la revolución digital. Se prevé que, en 2050, el 70 por ciento de la población mundial vivirá en ciudades [FP+12]. Uno de los grandes retos que prácticamente todos los países van a enfrentar en este siglo es la planificación, administración y gobernanza de las ciudades de forma sostenible, maximizando las oportunidades económicas y minimizando los daños medioambientales. Es necesario utilizar mejor los recursos públicos y explotar los activos naturales de forma consciente y responsable. Transformar “ciudades tradicionales” en Ciudades Inteligentes, o *Smart Cities* es una demanda cada vez más importante.

A medida que el paradigma de Internet de las cosas (en inglés *Internet of Things - IoT*) se va extendiendo, las tecnologías de la información y la comunicación (TIC) tienen un cometido fundamental que desempeñar para incrementar la eficiencia en todos los sectores industriales y permitir innovaciones tales como los sistemas de transporte inteligentes y la gestión "inteligente" del agua, de la energía y de los residuos, entre otros.

En un futuro cercano con IoT los objetos de la vida cotidiana estarán equipados con microcontroladores, transreceptores para comunicaciones digitales y varios protocolos adecuados que les permita comunicarse unos con otros y con los usuarios, convirtiéndose en una parte integral de la Internet [AIM10]. El concepto IoT, por lo tanto, tiene como objetivo hacer que Internet sea aún más inmersivo y omnipresente. Además, al permitir un fácil acceso y la interacción con una amplia variedad de dispositivos

tales como, electrodomésticos, cámaras de vigilancia, sensores, actuadores, pantallas, vehículos, etc., la IoT fomenta el desarrollo de una serie de aplicaciones que hacen uso de la cantidad potencialmente enorme y variedad de datos generados por tales objetos para proporcionar nuevos servicios a los ciudadanos, empresas y a las organizaciones con aplicación en diferentes ámbitos ([BC+13],[SD+16],[GS+17] y [Kun14]).

Una red de sensores inalámbricos (en inglés, *Wireless Sensor Network, WSN*) es un conjunto de nodos sensores y actuadores desplegados sobre la región de actuación. Cada nodo sensor consiste de una o varias capacidades de procesamiento, memoria, transceptor de comunicación, una fuente de energía, y funciones específicas de sensado o actuación. Los sistemas SCADA (Supervisión, Control y Adquisición de Datos) se aplican a la administración de la recolección de información de los procesos de controles industriales, tales como las redes eléctricas, redes de distribución de agua, transporte, fabricación, etc.).

Los sensores son dispositivos económicos, de bajo consumo de energía con recursos limitados [AS+02]. Una WSN se compone de un gran número de nodos de sensores con capacidad limitada de computación, almacenamiento y comunicación. Las WSN, representan una de las áreas de investigación más prometedoras debido al completísimo campo de aplicación y el desarrollado de nuevas tecnologías de controladores electrónicos o electromecánicos, y los avances en la tecnología de comunicación inalámbrica [Cal04]. El diseño de una WSN debe hacer frente a la ocurrencia de defectos de alguna o múltiples

componentes en la red y por lo tanto es necesario la aplicación de técnicas robustas que puedan tolerar los fallos, sin resentir el buen funcionamiento de la red. Los problemas y las necesidades de mecanismos de mejora de la fiabilidad dependen de los recursos y aplicaciones disponibles para el que se aplica la WSN [VSS13].

La expansión de *Big Data* y la evolución de las tecnologías de IoT juegan un rol importante en la factibilidad de las iniciativas para ciudades inteligentes. *Big Data* ofrece potencial para que las ciudades obtengan conocimiento de valor de grandes cantidades de información colectada de varias fuentes y la IoT permite la integración de sensores, identificación de radiofrecuencias, y *bluetooth* en el entorno del mundo real utilizando servicios en red. La combinación del IoT y Big Data es un área de investigación poco explorada que ha traído nuevos e interesantes retos para alcanzar el objetivo de futuras ciudades inteligentes [RA+16], [ZL+15], [BD+16], [JJ+15].

El aumento significativo de dispositivos y sensores conectados han hecho factible la visión de vivir en un entorno inteligente. Se ha cambiado del ámbito de la computación de escritorio tradicional a una computación cada vez más sofisticada [GB+13]. Recientemente se han introducido varias aplicaciones de un entorno inteligente, incluidas las casas inteligentes [CDN11], las redes inteligentes [CS+09], el transporte inteligente [AJ09], el cuidado de la salud inteligente [Dem13], y las ciudades inteligentes [CND11], [CN+12], debido al crecimiento de la población urbana y la rápida urbanización.

El fenómeno de Big Data desde hace tiempo se ha caracterizado por el volumen, la velocidad y una variedad de tipos de datos que se incrementa día a día. Big Data ofrece el potencial para que la ciudad obtenga información valiosa de una cantidad considerable de datos recogidos a través de diversas fuentes. Ciertamente, las particularidades de estos datos incluyen características no estructuradas en comparación con grandes datos recogidos por otros medios [CML14]. Por ejemplo, identificar y obtener información valiosa de grandes cantidades de datos meteorológicos puede ser extremadamente beneficioso en términos de desarrollo agrícola. Además, el análisis de los datos meteorológicos puede ayudar a informar a la gente con antelación sobre las posibles condiciones peligrosas (información sobre inundaciones, clima extremadamente caluroso, sequía, etc.).

Los gobiernos han comenzado a adoptar ideas inteligentes para mejorar el nivel de vida de sus ciudadanos e implementar grandes aplicaciones de datos. La ciudad inteligente aprovecha las tecnologías emergentes, como las redes de sensores inalámbricos (WSN), para reducir el costo y el consumo de recursos. Sin embargo, la aplicación del análisis de grandes datos en el entorno inteligente se mantiene aún en las etapas iniciales

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta sección se describe la línea de investigación que se lleva a cabo en el proyecto.

Se puede decir que una ciudad inteligente es la que realiza actividades en al menos una de las seis áreas de acción inteligente: Economía inteligente, Sociedad inteligente, Gobernanza

inteligente, Movilidad inteligente, Medioambiente inteligente, Modo de vida inteligente.

Algunas aplicaciones que se pueden definir en estas áreas y que son de interés en este trabajo de investigación son: (a) Movilidad: rutas dinámicas personalizadas; seguridad urbana y (b) Medioambiente: gestión inteligente de la basura; energía inteligente.

Movilidad inteligente: La movilidad es otro gran reto que deben afrontar las grandes ciudades por el importante crecimiento en la cantidad de vehículos. Esos vehículos circulan por las vías públicas y contribuyen al aumento de los congestionamientos, los accidentes graves, las emisiones de gases contaminantes y también de los gases de efecto invernadero.

Por otro lado, la gestión adecuada de los residuos urbanos es otro tema de creciente preocupación para los agentes públicos, con impactos directos sobre la salud, el medio ambiente y la calidad de vida de la población. Las ciudades son grandes emisores de gas metano (CH₄), con un potencial de calentamiento global 21 veces mayor que el gas dióxido de carbono (CO₂). Según un estudio de la ONU, la actual generación de basura en el mundo maneja alrededor de 1.300 millones de toneladas/año, y se prevé que para 2025 llegará a 2.200 millones de toneladas/año. Los costos financieros y ambientales de administrar un volumen tan significativo de residuos son enormes. Además, los servicios de recolección de residuos trabajan en condiciones de incertidumbre, es decir, la cantidad de residuos que generarán los ciudadanos y cuándo se desecharán es desconocida, pero ciertamente muestran algunos patrones. Entonces, la recolección de información sobre la capacidad real de

llenado de los contenedores en tiempo real es muy importante.

Para abordar estas áreas se deberá usar información existente (datos abiertos), datos del presente (sensores) y predicciones. Los desafíos son varios e involucran la combinación de técnicas de optimización, aprendizaje de máquina y análisis de datos para obtener soluciones usables y perdurables en ciudades inteligentes y sostenibles.

3. RESULTADOS OBTENIDOS/ESPERADOS

La solución de problemas a través de los dos ejes de desarrollo en ciudades inteligentes permitirá hacer prototipos reales, aprender más y producir un mayor impacto en Ciencia e Industria. Una ciudad inteligente es un nuevo dominio, la visión holística es obligatoria. El impacto será mayor al considerar: ciencia e industria. El desafío a las técnicas inteligentes será más alto, así como también el conocimiento que se construya. Dada la amplia visión se espera generar oportunidades para nuevos proyectos y socios.

Además, se espera realizar la publicación de los resultados del proyecto en revistas de alto impacto y conferencias de los ámbitos multidisciplinarios.

Finalmente, se pretende colaborar con los gobiernos locales para la implementación de políticas y acciones inteligentes y sostenibles que impacten en la calidad de vida de los ciudadanos.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto de investigación proporcionará un marco propicio para la iniciación y/o finalización de estudios de

posgrado de los integrantes docentes. De igual forma, será un ámbito adecuado para la realización de tesis de grado. Actualmente dos integrantes están desarrollando sus tesis de Maestría y dos sus tesis de Doctorado. Además, se cuenta con un becario de grado y un becario de posgrado.

5. BIBLIOGRAFIA

- [AJ09] Adeli, H., & Jiang, X. (2009). Intelligent infrastructure: neural networks, wavelets, and chaos theory for intelligent transportation systems and smart structures. Crc Press.
- [AS+02] Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). A survey on sensor networks. *IEEE Communications magazine*, 40(8), 102-114.
- [BC+13] Bellavista, P., Cardone, G., Corradi, A., & Foschini, L. (2013). Convergence of MANET and WSN in IoT urban scenarios. *IEEE Sensors Journal*, 13(10), 3558-3567.
- [BD+16] Botta, A., De Donato, W., Persico, V., & Pescapé, A. (2016). Integration of cloud computing and internet of things: a survey. *Future Generation Computer Systems*, 56, 684-700.
- [Cal04] Callaway Jr, E. H. (2003). *Wireless sensor networks: architectures and protocols*. CRC press.
- [CDN11] Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart cities in Europe. *Journal of urban technology*, 18(2), 65-82.
- [CML14] Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, 19(2), 171-209.
- [CN+12] Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., ... & Scholl, H. J. (2012, January). Understanding smart cities: An integrative framework. In *System Science (HICSS), 2012 45th Hawaii International Conference on* (pp. 2289-2297). IEEE.
- [CS+09] Chen, S. Y., Song, S. F., Li, L. X., & Shen, J. (2009). Survey on smart grid technology [J]. *Power System Technology*, 8, 1-7.
- [Dem13] Demirkan, H. (2013). A smart healthcare systems framework. *It Professional*, 15(5), 38-45.
- [FP+12] Fazio, M., Paone, M., Puliafito, A., & Villari, M. (2012, July). Heterogeneous sensors become homogeneous things in smart cities.
- [GB+13] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future generation computer systems*, 29(7), 1645-1660.
- [GS+17] Gharaibeh, A., Salahuddin, M. A., Hussini, S. J., Khreishah, A., Khalil, I., Guizani, M., & Al-Fuqaha, A. (2017). Smart Cities: A Survey on Data Management, Security and Enabling Technologies. *IEEE Communications Surveys & Tutorials*. 775-780). IEEE.
- [JJ+15] Jaradat, M., Jarrah, M., Bouselham, A., Jararweh, Y., & Al-Ayyoub, M. (2015). The internet of energy: Smart sensor networks and big data management for smart grid. *Procedia Computer Science*, 56, 592-597.
- [Kun14] Kunzmann, K. R. (2014). Smart cities: a new paradigm of urban development. *Crios*, 4(1), 9-20.
- [SD+16] Sarasola, B., Doerner, K. F., Schmid, V., & Alba, E. (2016). Variable neighborhood search for the stochastic and dynamic vehicle routing problem. *Annals of Operations Research*, 236(2), 425-461.
- [RA+16] Rathore, M. M., Ahmad, A., Paul, A., & Rho, S. (2016). Urban planning and building smart cities based on the internet of things using big data analytics. *Computer Networks*, 101, 63-80.

[VSS13] Venkatesan, L., Shanmugavel, S., & Subramaniam, C. (2013). A survey on modeling and enhancing reliability of wireless sensor network. *Wireless Sensor Network*, 5(03), 41.

[ZL+15] Zhu, C., Leung, V. C., Shu, L., & Ngai, E. C. H. (2015). Green Internet of Things for smart world. *IEEE Access*, 3, 2151-2162.

Metaheurísticas aplicadas de problemas de scheduling con restricciones

Pandolfi D.¹, Villagra A.¹, Leguizamón G.^{1,2}, Orozco S.¹, Rasjido J.¹, Varas V.¹, Seron N.
{dpandolfi, avillagra,sorozco,jrsjido,vvaras,nseron,}@uaco.unpa.edu.ar, legui@unsl.edu.ar

¹ Laboratorio de Tecnologías Emergentes (LabTEm)
Instituto de Tecnología Aplicada - Unidad Académica Caleta Olivia
Universidad Nacional de la Patagonia Austral

² Laboratorio de Investigación y Desarrollo en Inteligencia Computacional
(LIDIC) Departamento de Informática - Universidad Nacional de San Luis

Resumen

La mayoría de los problemas de optimización incluyen restricciones de cierta clase, que constituyen grandes desafíos a la resolución de problemas de optimización. Las restricciones son límites impuestos a las variables de decisión y en general las restricciones son una parte integral de la formulación de cualquier problema.

Los problemas de *Scheduling* consisten en la asignación de tareas a recursos limitados donde ciertos objetivos deben optimizarse y varias restricciones deben cumplirse. La mayoría de los problemas del mundo real tienen varios objetivos que tratamos de optimizar al mismo tiempo. Particularmente, la planificación de las actividades en un yacimiento que requiere de un proceso altamente complejo e implica un número considerable de actividades sujetas a restricciones. Las actividades en los pozos WAS (en inglés, *Well Activity Scheduling*) se ocupa de la coordinación

para formar así cronogramas de actividades que deben cumplir un conjunto de restricciones.

Este proyecto propone desarrollar algoritmos metaheurísticos, que incorporen heurísticas y reglas de despacho que sean competitivas con los algoritmos del estado del arte. Los enfoques propuestos tomarán como base metaheurísticas tales como cGA (*cellular Genetic Algorithm*) y técnicas de manejo de restricciones tales como funciones de penalidad y algoritmos de reparación.

Palabras clave: Metaheurísticas, Scheduling, Manejo de Restricciones

Contexto

La línea de trabajo se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEm), Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Caleta Olivia Universidad Nacional de la Patagonia Austral, en el marco del proyecto “Inteligencia Computacional aplicada a la optimización

multiojetivo de problemas de *scheduling* con restricciones”.

Introducción

Los problemas de planificación (*Planing Problems*) son una sub-clase de problemas de optimización combinatoria en campos tales operaciones de producción y despacho en la industria manufacturera y extractiva. Particularmente los problemas de secuenciamiento de tareas conocidos como *Scheduling Problems* tienen su aplicación en la industria y por lo tanto un fuerte impacto económico y social, donde el objetivo principal de los investigadores es la reducción de los costos de producción en la industria [Leu04].

Los problemas de *Scheduling* incluyen la combinación de recursos, tareas, objetivos y restricciones, donde un incremento en el tamaño del problema produce un incremento exponencial del espacio de soluciones. Muchos de los problemas de *Scheduling* son computacionalmente complejos y el tiempo requerido para calcular una solución óptima se incrementa con el tamaño del problema. Además, se ha demostrado, que muchos problemas de *Scheduling* pertenecen a la clase de NP-Hard [Bru04]

Las Metaheurísticas (MHs) son métodos que integran de diversas maneras, procedimientos de mejora local y estrategias de alto nivel para crear un proceso capaz de escapar de óptimos locales y realizar una búsqueda robusta en el espacio de búsqueda. En su evolución,

estos métodos han incorporado diferentes estrategias para evitar la convergencia a óptimos locales, especialmente en espacios de búsqueda complejos. En otras palabras, las MHs proveen de un marco general que permite crear nuevos híbridos a través de la combinación de conceptos derivados de: heurísticas clásicas, inteligencia artificial, evolución biológica, sistemas naturales, mecánica estadística, etc.

La complejidad de muchos problemas optimización discreta del mundo real está asociada con grandes espacios de búsqueda, demandas de rendimiento de tiempo real y ambientes dinámicos que no pueden ser resueltos por métodos exactos en tiempo razonable. Esto ha promovido en la comunidad científica el desarrollo de nuevos enfoques conocidos como metaheurísticas [CD09].

Los cGAs son una subclase de metaheurísticas con una población estructurada espacialmente, es decir, los individuos de la población pueden aparearse solo con sus vecinos. Los cGAs, se diseñaron inicialmente para trabajar en máquinas paralelas, formadas por muchos procesadores que ejecutaban simultáneamente las mismas instrucciones sobre diferentes datos (máquinas SIMD - El primer modelo de cGA fue propuesto por Robertson en 1987 [Rob87] y fueron inicialmente desarrollados para trabajar en máquinas paralelas, y más tarde se fueron adaptando para funcionar también en máquinas de procesamiento secuencial.

La mayoría de los problemas de optimización incluyen restricciones de cierta clase, que constituyen grandes desafíos a la resolución de problemas de optimización. Las restricciones son límites impuestos a las variables de decisión y en general las restricciones son una parte integral de la formulación de cualquier problema [DR90]. El manejo de restricciones en algoritmos evolutivos puede abordarse de distintas técnicas y un amplio estudio del estado del arte es propuesto por Coello Coello [Coe02].

La forma más común de incorporar restricciones en un algoritmo evolutivo ha sido a través de funciones de penalidad [Sch95]. La idea de este método es transformar un problema de optimización con restricciones en un uno sin restricciones agregando (o sustrayendo) un cierto valor a la función objetivo basado en la cantidad de violación la restricción presentada en una solución. Las técnicas basadas en algoritmos de reparación mapean una solución desde el espacio no factible al espacio factible. Los procesos de reparación que modifican genéticamente la solución no factible son conocidos como Lamarckianos y transforman una solución no factible en factible la que es evolucionada por el algoritmo. Un enfoque menos destructivo de la solución no factible, permite combinar el aprendizaje y la evolución, estrategia conocida como Baldwiniana. En este último enfoque las soluciones son reparadas solo para su evaluación. Estudios analíticos y empíricos indican que esta técnica reduce la velocidad de convergencia del algoritmo evolutivo y

permite converger a óptimo globales [WG+94].

Líneas de investigación y desarrollo

Las empresas petroleras realizan visitas de mantenimiento preventivo a cada una de sus locaciones petroleras (pozos productores, inyectores, baterías y colectores). Un yacimiento está formado por bloques y a su vez éste por baterías. Cada batería está formada por pozos de producción que son en promedio entre 15 y 20. Cada pozo tiene diferente nivel de producción que es conocido a priori y varía en el tiempo. La producción del pozo define la categoría y la cantidad de veces que debe visitarse al mes. Los pozos no pueden ser visitados más de una vez al día y dependiendo del tipo de pozo existen ciertas tareas que se deben realizar. Cada tarea tiene asignado un determinado equipamiento necesario, una frecuencia de realización y un tiempo aproximado de su duración.

Este proyecto propone desarrollar un conjunto metaheurísticas basadas en un Inteligencia Computacional que proporcionen información detallada para decisiones a corto y largo plazo, mono-criterio, multi-criterio. Esta herramienta permitirá determinar el orden óptimo de perforación, hora y lugar, tipos de pozos y por lo tanto el número de nuevos pozos inyectores y productores. Además, proporcionará planes de mantenimiento correctivos y preventivos. Todos los planes se deben cumplir las restricciones de actividades y recursos dispuestos y para los criterios de optimización establecidos.

Resultados obtenidos/esperados

En esta línea de trabajo se pretende:

a) Estudiar estado del arte para algoritmos metaheurísticos, heurísticas y reglas de despacho para distintos problemas de scheduling (máquina única, flow shop, job shop), optimización multiojetivo y mecanismos para el tratamiento de restricciones.

b) Implementar motores de planificación basados en metaheurísticas con restricciones.

c) Estudiar los modelos de procesos para la planificación de actividades del desarrollo de reservorios de gas y petróleo. (actividades, restricciones).

Formación de recursos humanos

Esta línea de investigación proporcionará un marco propicio para la iniciación y/o finalización de estudios de posgrado de los integrantes docentes. De igual forma, será un ámbito adecuado para la realización de tesis de grado. En ese sentido, dos integrantes de este proyecto de investigación está desarrollando su Tesis de Maestría en temáticas afines y un integrante está desarrollando su Tesis de doctorado. Además, se cuenta con un becario alumno de la carrera de Ingeniería en Sistemas.

Referencias

- [Bru04] Brucker. P. (2004) Scheduling Algorithms. Springer-Verlag 3rd ed.
- [CD09] Chiong R. and Dhakal S., (2009). Natural intelligence for scheduling, planning and packing problems. Springer, Berlin.
- [Coe02] Coello Coello C., (4 January 2002), Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art, Computer Methods in Applied Mechanics and Engineering, Vol. 191, No. 11- 12, pp. 1245-128.
- [DR90] Dhar, V., and Ranganathan, N. (1990). Integer programming vs. expert systems: an experimental comparison. Communications of the ACM, 33(3), 323-336.
- [6] Ehrgott M. (2005), Multicriteria Optimization. Springer, Berlin, second edition, 2005. ISBN 3-540-21398-8.
- [7] Framinan J., Gupta J., and Leisten R. (2004). A review and classification of heuristics for permutation flow-shop scheduling with makespan objective. Operational Research Society, 55(12):1243-55.
- [Leu04] Leung J. (2004) Handbook of Scheduling: Algorithms, Models and Performance Analysis. Computer and Information Sciences Series. Chapman and Hall.
- [Rob87] G. Robertson. Parallel implementation of genetic algorithms in a classier system. In Second International Conference on Genetic Algorithms (ICGA), pages 140{147, 1987.
- [Sch95] Schwefel P. (1995), Evolution and Optimum Seeking. John Wiley & Sons, New York.
- [WG+94] Whitley, D., Gordon, S., and Mathias, K.: Lamarckian evolution, the Baldwin effect and function optimization, Parallel Problem Solving from Nature (1994).

ANÁLISIS AUTOMÁTICO DE GRANDES VOLÚMENES DE DATOS EN REDES SOCIALES MEDIANTE MINERÍA DE TEXTOS COMBINADO CON ALGORITMOS INTELIGENTES

Juan Pablo Tessore ^{1,2}, Leonardo Esnaola ², Claudia Russo ^{2,3}, Hugo Ramón ^{2,3}, Sabrina Pompei ²

Instituto de Investigación y Transferencia en Tecnología (ITT)⁴
Comisión de Investigaciones Científicas (CIC)
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires
(UNNOBA)

Sarmiento y Newbery, 236-4636945/44

{juanpablo.tessore, leonardo.esnaola, claudia.russo, hugo.ramon, sabrina.pompei}@itt.unnoba.edu.ar

RESUMEN

El presente trabajo propone construir un clasificador automático de opiniones, que permitirá realizar análisis automáticos a bajo costo del juicio de los consumidores acerca de productos o servicios. Dicho clasificador será entrenado a partir de los comentarios en lenguaje informal presente en redes sociales.

Para alcanzar el objetivo descripto, en primer lugar, se prevé construir una base de datos que reúna diversos fragmentos de texto en idioma español, incorporando los modismos propios de nuestra región.

En segundo lugar, a través de un proceso incremental de limpieza y normalización de cada fragmento de texto, que incluye actividades como la eliminación de hashtags, enlaces, emoticones, etc.; corrección

ortográfica; etiquetado sintáctico (también conocido como “Part Of Speech Tagging”, o simplemente “POS tagging”); desambiguación, entre otras.

Una vez realizada la recopilación y normalizado el contenido, se definirá un criterio de clasificación de dichos fragmentos, de manera de establecer clases que permitan agrupar los mismos según su afinidad, es decir a partir de características comunes.

Finalmente, a partir del diseño, desarrollo e implementación de un algoritmo inteligente se buscará determinar el grado de pertenencia a cada uno de los grupos definidos de cualquier texto arbitrario.

Palabras clave: Text mining, Big Data, Inteligencia artificial, Redes sociales.

1 Becario de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

2 Docente Investigador en el Instituto de Investigación y Transferencia en Tecnología (ITT) / Escuela de Tecnología / UNNOBA

3 Investigador Asociado Adjunto sin director a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

4 Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

CONTEXTO

Esta línea de investigación forma parte del proyecto “Tecnología y Aplicaciones de Sistemas de Software: Calidad e Innovación en procesos, productos y servicios” aprobado por la Secretaría de Investigación, Desarrollo y Transferencia de la UNNOBA en el marco de la convocatoria a Subsidios de Investigación Bianuales (SIB2017). A su vez se enmarca en el contexto de un plan de trabajo aprobado por la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires y por la Secretaría de Investigación de la UNNOBA en el marco de la convocatoria “Becas de Estudio Cofinanciadas 2015 CIC Universidades del interior bonaerense”.

El proyecto se desarrolla en el Instituto de Investigación en Tecnologías y Transferencia (ITT) dependiente de la mencionada Secretaría, y se trabaja en conjunto con la Escuela de Tecnología de la UNNOBA.

El equipo está constituido por docentes e investigadores pertenecientes al ITT y a otros Institutos de Investigación, así como también, estudiantes de las carreras de Informática de la Escuela de Tecnología de la UNNOBA.

1. INTRODUCCIÓN

Desde que las primeras computadoras programables fueron concebidas, las personas se preguntaron si tendrían la capacidad de pensar, de aprender y de convertirse en “máquinas inteligentes”.

El campo de la ciencia que se encarga de resolver este interrogante se denomina inteligencia artificial. Se trata de un área multidisciplinaria, que a través de ciencias como las ciencias de la computación, la

matemática, la lógica y la filosofía, estudia la creación y diseño de sistemas capaces de resolver problemas cotidianos por sí mismos, utilizando como paradigma la inteligencia humana [1]. Para que una máquina pueda comportarse de manera inteligente debería ser capaz de resolver problemas de la manera en que lo hacen los humanos, es decir, en base a la experiencia y el conocimiento [2]. Esto implica que debería ser capaz de modificar su comportamiento en base a cuán precisos son los resultados obtenidos comparados con los esperados.

En este sentido podemos encontrar tres grandes grupos de algoritmos de *Machine Learning* [3]:

- Algoritmos supervisados: estos algoritmos utilizan un conjunto de datos de entrenamiento etiquetados (preclasificados), los cuales procesan para realizar predicciones sobre los mismos, corrigiéndolas cuando son incorrectas. El proceso de entrenamiento continúa hasta que el modelo alcanza un nivel deseado de precisión.
- Algoritmos semi-supervisados: combinan tanto datos etiquetados como no etiquetados para generar una función deseada o clasificador. Este tipo de modelos deben aprender las estructuras para organizar los datos así como también realizar predicciones.
- Algoritmos no supervisados: El conjunto de datos no se encuentra etiquetado y no se tiene un resultado conocido. Por ello deben deducir las estructuras presentes en los datos de entrada, lo puede conseguir a través de un proceso matemático para reducir la redundancia sistemáticamente u organizando los datos por similitud.

Dentro de esta clasificación podemos además encontrar un gran número de algoritmos específicos con diferentes características para el tratamiento de los datos. Entre los más relevantes encontramos:

- *Deep Learning* (DL): consiste en la utilización de algoritmos para hacer representaciones abstractas de la información y facilitar el aprendizaje automático [4].
- *Active Learning* (AL): es un caso especial de aprendizaje semi-supervisado donde el algoritmo de aprendizaje puede interactuar con un usuario u otra fuente de información para obtener los resultados deseados [5].
- *Support Vector Machines* (SVM): busca la maximización de la distancia entre la recta o el plano y las muestras que se encuentran a un lado u otro. En el caso que las muestras no sean linealmente separables se utiliza una transformación llamada *kernel* [6] [7].

Una de las principales áreas de la inteligencia artificial es el procesamiento del lenguaje natural (PLN) o minería de textos. Esta área se encarga de desarrollar algoritmos que permitan extraer información relevante a partir de diversos contenidos en forma de texto. Con el auge de los contenidos sociales, la generación de este tipo de contenidos ha crecido en forma exponencial, esto último crea la oportunidad de aplicar algoritmos de minería de textos para extraer patrones significativos [8].

Según [9], de entre las técnicas de *Machine Learning* mencionadas, SVM es la más comúnmente utilizada para el análisis automático de textos, esto es así principalmente porque esta técnica es más apropiada para resolver problemas con una gran cantidad de

dimensiones. Dentro de las tareas de minería de textos que pueden realizarse con esta técnica, podemos encontrar: clasificación de subjetividad; determinación de polaridad; resolución de ambigüedades; extracción de palabras de opinión y/o aspectos; etc.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La presente investigación se encuadra dentro de los ejes “Gestión de la innovación” y “Cómputo Ubicuo”. En ese sentido, los algoritmos basados en *Machine Learning* posibilitan, a través de su capacidad de aprendizaje y de su comportamiento inteligente de manera automática, el desarrollo de sistemas de cómputo ubicuo los cuales permiten, en última instancia, el monitoreo (y cambio) de la conducta humana y del ambiente donde esta se desarrolla.

Para la efectiva implementación de algoritmos de *Machine Learning*, se deberán abarcar las siguientes cuestiones:

- Obtención de un conjunto de datos, en formato de texto, suficientemente representativo para la problemática que se desea abordar.
- Pre procesamiento de las señales para lograr su normalización y adecuación.
- Determinación de categorías que permitan clasificar los textos recopilados según características comunes.
- Análisis y selección de los distintos algoritmos de minería de textos y *Machine Learning* que permitan clasificar los textos según las categorías definidas en el inciso anterior.

- Evaluación de fiabilidad y desempeño de las diferentes técnicas aplicadas.

3. RESULTADOS OBTENIDOS / ESPERADOS

Se espera que la presente línea de I/D permite adquirir conocimientos específicos sobre las diferentes técnicas de minería de textos y *Machine Learning*, con el propósito de desarrollar modelos capaces de predecir y clasificar los contenidos involucrados en la problemática que se intenta resolver, obteniendo un comportamiento inteligente de manera automática.

También se prevé la aplicación de *Machine Learning* y minería de textos en el análisis del texto redes sociales, con la finalidad de encontrar patrones dentro de esos datos que permitan predecir comportamientos futuros en ámbitos específicos.

Se espera como resultado final de esta línea de investigación, la creación de un producto transferible que permita llevar a cabo estudios sobre bienes y servicios, ofrecidos por personas y organizaciones pertenecientes a los sectores público y privado, detectando las opiniones manifestadas indirectamente por los comentarios de sus consumidores, sin la necesidad de destinar cuantiosos recursos a un análisis pormenorizado de los mismos.

Así mismo, se busca generar informes técnicos en base al trabajo realizado, en donde se registren los avances, el grado de implementación y los resultados obtenidos. Como así también difundir y transferir los resultados y logros alcanzados mediante la presentación y participación en diferentes congresos, jornadas y workshops de carácter

nacional e internacional vinculados a la temática de estudio.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de I/D se han obtenido y se encuentran desarrollando actualmente una Beca de Estudio Cofinanciada otorgada por la Comisión de Investigaciones Científicas (CIC) y la UNNOBA. Asimismo se espera desarrollar una tesis doctoral y dos tesinas de grado, dirigidas por miembros de este proyecto.

5. BIBLIOGRAFÍA

- [1] Assessment of the Commercial Applicability of Artificial Intelligence in Electronic Businesses. Thomas Kramer. Diplom.de. 2002.
- [2] Data Classification Algorithms and Applications, Charu C. Aggarwal, CRC Press, 2015.
- [3] Machine Learning An Algorithmic Perspective Second Edition, Stephen Marsland, CRC Press, 2015.
- [4] A Deep Learning. Book in preparation for MIT Press. Bengio, Y., Goodfellow, I. and Courville, USA, 2015.
- [5] Active Learning Literature Survey, Settles Burr, Computer Sciences Technical Report 1648. University of Wisconsin–Madison, 2014.
- [6] A Tutorial on Support Vector Machines for Pattern Recognition, Christopher J.C. Burges, Kluwer Academic Publishers, 1998.
- [7] Top 10 algorithms in data mining,

- Xindong Wu et al. Knowledge and Information Systems 2008.
- [8] Mastering Machine Learning with Python in Six Steps: A Practical Implementation Guide to Predictive Data Analytics Using Python, Manohar Swamynathan, APRESS, 2017.
- [9] A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Ravi Kumar et. al. IEEE Knowledge-Based Systems, 2015.

Integración de bases de creencias manteniendo coherencia y consistencia con criterio unificado de credibilidad

Cristhian A.D. Deagustini^{1,2}, Cristian D. Pacifico², Juan C.L. Teze^{1,2}, M. Vanina Martinez¹, Marcelo A. Falappa¹

¹Laboratorio de I+D en Inteligencia Artificial, Instituto de Ciencias e Ingeniería de la Computación (ICIC), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad Nacional del Sur, Alem 1253, (B8000CPB) Bahía Blanca, Arg.

²Área de Agentes y Sistemas Inteligentes, Fac. Cs Administración, Universidad Nacional de Entre Ríos, Tavella 1424, (E3202KAC) Concordia, Arg
cadd@cs.uns.edu.ar; cripac@fcad.uner.edu.ar; jcarlt02@gmail.com, {mvm,mfalappa}@cs.uns.edu.ar

RESUMEN

Esta línea de I+D propone desarrollar *frameworks* para la integración de bases de conocimientos provenientes de diferentes fuentes. Estos *frameworks* modelan escenarios de Revisión de Creencias en entornos de múltiple agentes, con mecanismos para solucionar conflictos de inconsistencia e incoherencia recurrentes al integrar varios repositorios masivos locales y de gestión diferente. Se pretende que el proceso de integración, logre consensuar un criterio unificado de credibilidad que refleja la confianza global en cada fuente; y posteriormente, efectivice la integración como una vista unificada de la información de los repositorios locales, basándose en el orden de credibilidad establecido.

Palabras clave: revisión de creencias, bases de credibilidad, sistemas multiagentes.

CONTEXTO

La presente investigación surge en el marco del proyecto **PID-UNER 7053** “Integración de bases de creencias manteniendo coherencia y consistencia: teoría y aplicaciones”, que es desarrollado en la Fac. Cs. de la Administración de la UNER y tiene integrantes de dicha Universidad y del Inst. de Ciencias e Ingeniería de la Computación de UNS-CONICET.

1. INTRODUCCIÓN

En sistemas que manejan e intercambian grandes volúmenes de información es necesario contar con métodos automáticos que solucionen los conflictos que aparecen cuando el conocimiento evoluciona y se integra con otras fuentes de información; más aún cuando esas porciones de conocimiento provienen de diferentes fuentes o dominios. Actualmente, es interesante poder representar tales Bases de Conocimiento como *ontologías*. En particular las *ontologías* son importantes no sólo por el conocimiento que expresan, sino también por las restricciones semánticas que pueden aplicarse. Si de expresar ontologías se trata, se destaca el lenguaje **Datalog±** [1]. Desde el punto de vista de expresividad, Datalog± permite construir representaciones de conocimiento basadas en reglas modulares que son sintácticamente fragmentos de Lógica de Primer Orden. Datalog± como máquina de inferencia, tiene la propiedad de ser decidible permitiendo manejar grandes volúmenes de datos de entornos reales.

Al contar con ontologías, surge la necesidad de manejar dinámica de la base de conocimiento. Es decir, aceptar nueva información manteniendo la integridad y consistencia de la base; reconociendo que el manejo de conflictos es un problema que debe ser atacado [2, 3], especialmente cuando se integra conocimiento proveniente de diferentes fuentes [4, 5].

El más conocido de los conflictos es la **inconsistencia** de la información. Este concepto se refiere a teorías tales que es

imposible encontrarles un *modelo* (es decir: una *interpretación* que haga verdaderas a todas sus fórmulas). En términos sintácticos, un conjunto de fórmulas es *consistente* si y sólo si, dada una fórmula **A**, no es posible deducir tanto **A** como $\neg\mathbf{A}$. Adicionalmente, en entornos ontológicos se encuentra otro tipo de conflicto denominado *incoherencia* [6, 7]. Este fenómeno surge cuando una porción de conocimiento, expresado por un conjunto de reglas, no pueda ser aplicado sin generar problemas de consistencia; es decir, las reglas no puede ser aplicadas sin violar inevitablemente alguna de las restricciones impuestas al conocimiento, haciéndolas por lo tanto insatisfacibles. Intuitivamente se puede ver a la incoherencia como una inconsistencia *latente* o *potencial*. Consecuentemente, si este conjunto de reglas *incoherente* es considerado junto con hechos relevantes (hechos que activen las reglas en cuestión) entonces la violación es inevitable, generando *inconsistencia*. En estos entornos, una posibilidad para tratar estos conflictos y recuperar la coherencia y la consistencia implica modificar la información contenida en la ontología de forma tal de recuperar la consistencia –y coherencia- de la misma.

El área de I+D *Revisión de Creencias* [8, 9] (*Belief Revision –BR-*) ha realizado significativos aportes a la resolución de tales conflictos. Los marcos de trabajos de BR proponen operaciones para la dinámica de una base de conocimiento o conjunto de creencias. En el modelo fundacional AGM [10] de BR, dado conjunto de creencias **K** y una sentencia α , se establece operaciones de: **Expansión** $\mathbf{K}+\alpha$ (agrega una nueva α a **K**), **Contracción** $\mathbf{K}-\alpha$ (elimina toda posibilidad que de **K** se infiera α), y **Revisión** $\mathbf{K}*\alpha$ (asegura que de **K** se infiera α consistentemente, eliminando la posibilidad de inferir $\neg\alpha$). Otros trabajos posteriores desarrollaron otras operaciones, como ser **Consolidación** $\mathbf{K}!$ (*aplicable con **K** inconsistente, para volverla consistente*) [11], **Combinación** $\oplus(\Psi)=\mathbf{K}'$ (aplicable a un conjunto de bases de creencias $\Psi = \{\mathbf{K}_1, \mathbf{K}_2, \dots, \mathbf{K}_n\}$ y logra una única base \mathbf{K}' consistente respetando restricciones de integridad

impuestas) [5], e **Integración** $\Delta(\Psi) = \mathbf{K}'$ (que unifica diferentes bases teniendo en cuenta la fuente respectiva) [12].

Naturalmente, Revisión de Creencias ha evolucionado para abarcar los entornos multigente; de manera de poder definir y construir operaciones de cambio sobre las creencias de un agente individual, o las bases de creencias de una sociedad de agentes. Una síntesis de estos escenarios puede verse en [13]. Los ambientes (mono)agente, se identifican con *Revisión de la Creencia Individual (IBR)*, dónde se logra satisfacer los postulados AGM o frameworks equivalentes. En sistemas multiagentes, el primer caso es de *Revisión de Creencias de Múltiples Fuentes (MSBR)* dónde un agente individual recibe información de distintos agentes informantes y, eventualmente, agrega a su base de creencias conforme a un criterio de credibilidad sobre los otros informantes. Varios trabajos han sido presentados al respecto [14, 15]; y en particular [16, 17]. En el segundo escenario es la *Revisión de Creencias de Múltiples Agentes (MABR)*, dónde una sociedad de agentes deberá poder consensuar y unificar los criterios de credibilidad locales; y posteriormente lograr la integración de las bases locales en una única base de creencias de la sociedad de agentes; algunos avances se presentaron en [10, 11]. En estos dos últimos escenarios, el desafío centrar es poder integrar bases de conocimientos de diferentes fuentes.

La *integración (merging)* de múltiples fuentes de información es particularmente interesante para bases de datos distribuidas, sistemas multiagentes y para los sistemas distribuidos en general. En cuanto al manejo de la *inconsistencia* debida a datos contradictorios, se puede decir que la integración ha recibido menos esfuerzos que las otras operaciones. Se han propuesto muchos métodos para integración de creencias [19]–[21]; y también se han propuesto metodologías para compararlos entre sí, como destacado se puede citar a [22]. En el trabajo posterior de Konieczny y Pino Pérez [12], se proponen los postulados de varios operadores en un marco lógico;

además, se presentan algunas familias de operadores de integración que se estudian a la luz de estas propiedades.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Generalmente los grandes volúmenes de información están expresados en repositorios o bases de datos locales; y eventualmente necesitan integrarse en una federación de repositorios; obteniendo una vista unificada del conocimiento que expresan. Por tal motivo, esta propuesta se centra en la integración de tales bases locales para lograr una vista unificada global; interpretando a cada base de conocimiento local como una base de creencias de fuente (o agente) particular; y materializando la vista global unificada como un proceso de integración (*merging*) de las bases de creencias, que debe tener en cuenta la credibilidad particular de cada fuente (o agente); o logrando un criterio de credibilidad unificado.

Se sabe que Datalog \pm tiene reconocida capacidad para expresar información de bases de datos relacionales [1, 23]. Eventualmente se pueden hacer diferentes ontologías para representar bases de datos de diferentes fuentes o administraciones. Así, Datalog \pm ofrece un buen balance entre tratabilidad y expresividad para definir una infraestructura que permita la realización de consultas sobre grandes volúmenes de datos almacenados en repositorios federados.

Si bien nuestra propuesta no excluye el uso de otros formalismos, los trabajos ya publicado posibilitan continuar con líneas de acción ya planteadas [24]. En particular resulta interesante la definición de operadores de consolidación sobre bases de creencias Datalog \pm , presentada en [25]. Aquí se explora, como trabajo futuro, la posibilidad de expandir los postulados y representaciones para especificar operadores para integración de bases de creencias respetando un único orden de credibilidad.

Así, la hipótesis planteada es:

1. Que es posible representar una vista unificada de varias bases de conocimientos de diferentes fuentes mediante un *framework* basado en BR y especificando operadores de integración que manejen los potenciales conflictos de consistencia y coherencias.
2. Que tal escenario puede representarse como un entono de múltiples agentes, siendo cada uno el representante de una fuente determinada de información, con preferencias de credibilidad referidas al resto de los agentes.
3. Que es posible definir un criterio u orden de credibilidad unificado entre todos los agentes de información, y utilizar tal criterio para lograr la vista unificada.

3. RESULTADOS ESPERADOS

La validación de la hipótesis planteada supone la concreción de un *framework* que permitirá realizar integraciones de repositorios de administración separada; obteniendo una vista global, coherente y consistente, del conocimiento que estos repositorios manifiestan. Esta vista se logra por el consenso de criterios de credibilidad que asignan un valor de confianza a cada fuente. El propósito de cada vista, define la lógica de integración y afecta el criterio de credibilidad global. Se pueden constituir varias vistas, para varios propósitos. Actualmente los grandes repositorios de información están estructurados con esquemas relacionales de bases de datos. Determinados *frameworks*, como Datalog \pm , permiten expresar la información de tales repositorios en ontologías, sobre las cuales se pueden agregar lógicas de razonamiento, permitiendo inferir más información a partir de los hechos descriptos en las bases de datos. De concretar el *framework* de integración propuesto, es posible lograr un alto nivel de automatización en este proceso;

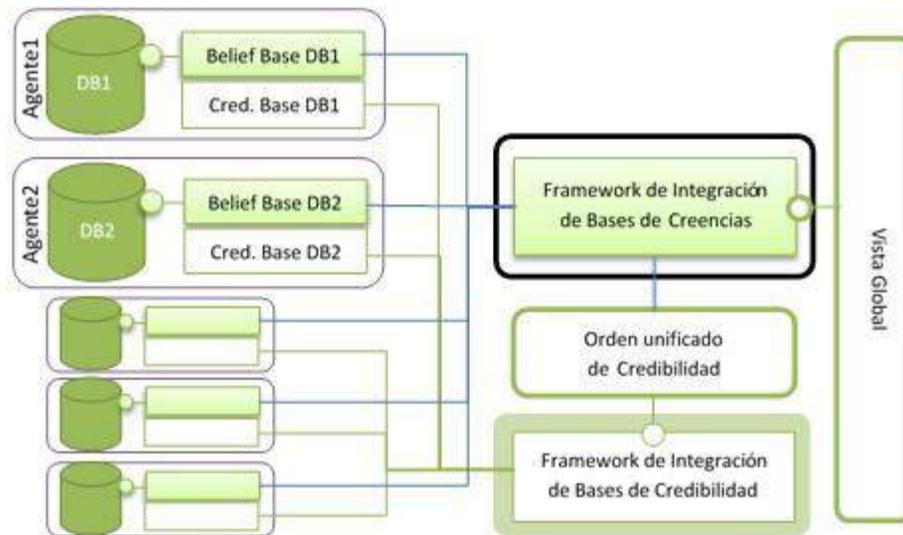


Figura 1: Estructura del *framework* propuesto

además de respetar la actualización continua que cada repositorio local pueda sufrir.

En una etapa posterior, se pueden desarrollar y construir aplicaciones informáticas que sirvan para lograr sintetizar grandes volúmenes de información distribuida, en varios campos de aplicación. En este punto es innovador la posibilidad de concretar un framework para ambientes de revisión de creencias de múltiples agentes (MABR), usando la expresividad de formalismos ontológicos como $Datalog_{\pm}$. La Fig. 1 resume la arquitectura propuesta.

Según las siguientes actividades, los resultados esperados son:

1. *Definición de operaciones de integración de bases de credibilidad*: se pretende definir los postulados, teorema de representación y construcciones. Se realizarán análisis comparativos.
2. *Definición de operadores de integración de bases de creencias basados en criterios de credibilidad unificados*: se pretende definir los postulados, teorema de representación y construcciones candidatas. En particular, también se definirá la estructura del estado epistémico y entradas epistémicas.
3. *Definición de operadores de integración de bases de creencias $Datalog_{\pm}$ basados en criterios de credibilidad unificados*: se espera lograr la definición de postulados, teorema de representación y

construcciones. En particular mediante relaciones de orden de credibilidad.

4. *Definición del marco de trabajo conceptual para integrar bases de creencias teniendo en cuenta la credibilidad de fuentes*: se espera definir un marco de trabajo para realizar integración de bases de creencias de fuentes diferentes, conformando escenarios del tipo MABR.

4. FORMACIÓN DE RR.HH.

La línea de I+D es llevada adelante por un equipo de integrantes de la FCAD/UNER y el ICIC/UNS. Esto permite que la experiencia de un grupo de investigación consolidado en el área sirva para el apuntalamiento y desarrollo de investigación del área "Agentes y Sistemas Inteligentes" de la FCAD/UNER, la cual cuenta con una estructura y desarrollo embrionario. Se desprende que esta sinergia favorece la formación de recursos humanos en la FCAD/UNER en cuarto nivel, especializados en el área de Sistemas de Razonamiento.

Este trabajo es el marco para 1(un) proyecto de tesis Doctoral y 1(un) proyecto de tesis de Maestría. Adicionalmente se prevé, para los detalles de implementación de herramientas, la realización de 1 (una) Tesina de grado.

5. BIBLIOGRAFÍA

- [1] A. Cali, G. Gottlob, y T. Lukasiewicz, «A general Datalog-based framework for tractable query answering over ontologies», *Web Semant. Sci. Serv. Agents World Wide Web*, vol. 14, pp. 57-83, 2012.
- [2] D. A. Bell, G. Qi, y W. Liu, «Approaches to Inconsistency Handling in Description-Logic Based Ontologies», en *OTM Workshops (2)*, 2007, pp. 1303-1311.
- [3] Z. Huang, F. van Harmelen, y A. ten Teije, «Reasoning with Inconsistent Ontologies», en *IJCAI*, 2005, pp. 454-459.
- [4] L. Amgoud y S. Kaci, «An argumentation framework for merging conflicting knowledge bases: The prioritized case», en *ECSQARU*, 2005, pp. 527-538.
- [5] C. Baral, S. Kraus, y J. Minker, «Combining Multiple Knowledge Bases», *IEEE Trans Knowl Data Eng.*, vol. 3, n.º 2, pp. 208-220, 1991.
- [6] G. Flouris, Z. Huang, J. Z. Pan, D. Plexousakis, y H. Wache, «Inconsistencies, Negations and Changes in Ontologies», en *AAAI*, 2006, pp. 1295-1300.
- [7] G. Qi y A. Hunter, «Measuring Incoherence in Description Logic-Based Ontologies», en *ISWC/ASWC*, 2007, pp. 381-394.
- [8] S. O. Hansson, «Belief Base Dynamics», Uppsala University, Department of Philosophy, Uppsala, Sweden, 1991.
- [9] P. Gärdenfors, *Belief revision*, vol. 29. Cambridge University Press, 2003.
- [10] C. Alchourrón, P. Gärdenfors, y D. Makinson, «On the Logic of Theory Change: Partial Meet Contraction and Revision Functions», *J. Symb. Log.*, vol. 50, n.º 2, pp. 510-530, 1985.
- [11] S. O. Hansson, *A Textbook of Belief Dynamics: Solutions to Exercises*. Norwell, MA, USA: Kluwer Academic Publishers, 2001.
- [12] S. Konieczny y R. P. Pérez, «Merging Information Under Constraints: A Logical Framework», *J. Log. Comput.*, vol. 12, n.º 5, pp. 773-808, 2002.
- [13] W. Liu y M.-A. Williams, «A Framework for Multi-Agent Belief Revision, Part I: The Role of Ontology», en *Advanced Topics in Artificial Intelligence, 12th Australian Joint Conference on Artificial Intelligence, AI '99, Sydney, Australia, December 6-10, 1999, Proceedings*, 1999, pp. 168-179.
- [14] S. Benferhat, D. Dubois, y H. Prade, «Argumentative inference in uncertain and inconsistent knowledge bases», *CoRR*, vol. abs/1303.1503, 2013.
- [15] A. F. Dragoni y P. Puliti, «Distributed Belief Revision versus Distributed Truth Maintenance», en *Sixth International Conference on Tools with Artificial Intelligence, ICTAI '94, New Orleans, USA, November 6-9, 1994*, 1994, pp. 499-505.
- [16] L. H. Tamargo, A. J. García, M. A. Falappa, y G. R. Simari, «On the revision of informant credibility orders», *Artif Intell.*, vol. 212, pp. 36-58, 2014.
- [17] L. H. Tamargo, A. J. García, M. A. Falappa, y G. R. Simari, «Consistency maintenance of plausible belief bases based on agents credibility», en *12th International Workshop on Non-Monotonic Reasoning (NMR)*, 2008, pp. 50-58.
- [18] L. H. Tamargo, A. J. García, M. A. Falappa, y G. R. Simari, «Modeling knowledge dynamics in multi-agent systems based on informants», *Knowl. Eng. Rev.*, vol. 27, n.º 1, pp. 87-114, 2012.
- [19] C. Baral, S. Kraus, J. Minker, y V. S. Subrahmanian, «Combining Knowledge Bases Consisting of First Order Theories», en *Methodologies for Intelligent Systems, 6th International Symposium, ISMIS '91, Charlotte, N.C., USA, October 16-19, 1991, Proceedings*, 1991, pp. 92-101.
- [20] L. Cholvy y A. Hunter, «Information Fusion in Logic: A Brief Overview», en *Qualitative and Quantitative Practical Reasoning, First International Joint Conference on Qualitative and Quantitative Practical Reasoning ECSQARU-FAPR '97, Bad Honnef, Germany, June 9-12, 1997, Proceedings*, 1997, pp. 86-95.
- [21] V. S. Subrahmanian, «Amalgamating knowledge bases», *ACM Trans Database Syst.*, vol. 19, n.º 2, pp. 291-331, jun. 1994.
- [22] S. Konieczny y R. P. Pérez, «Merging with Integrity Constraints», en *Symbolic and Quantitative Approaches to Reasoning and Uncertainty, European Conference, ECSQARU'99, London, UK, July 5-9, 1999, Proceedings*, 1999, pp. 233-244.
- [23] T. M. Connolly y C. E. Begg, *Database Systems: A Practical Approach to Design, Implementation and Management*, 5th ed. USA: Addison-Wesley Publishing Company, 2009.
- [24] C. A. D. Deagustini, M. V. Martínez, M. A. Falappa, y G. R. Simari, «Inconsistency resolution and global conflicts», en *Proc. of 21st European Conference on Artificial Intelligence (ECAI '14)*, 2014, pp. 991-992.
- [25] C. A. D. Deagustini, «Consolidación de ontologías Datalog+», Departamento de Ciencias e Ingeniería de la Computación. Universidad Nacional del Sur., 2016.

Optimización a gran escala usando metaheurísticas

Juan José Barbero, Martín Tamagusku, Eber Bezzone,
Fernando Sanz Troiani, Hugo Alfonso¹,
Carlos Bermudez, Gabriela Minetti¹, Carolina Salto¹
Laboratorio de Investigación en Sistemas Inteligentes (LISI)
Facultad de Ingeniería - Universidad Nacional de La Pampa
Calle 110 Esq. 9 (6360) General Pico - La Pampa - Rep. Argentina
Te. / Fax: (02302) 422780/422372, Int. 6302
e-mail: ¹{minettig, saltoc, alfonsoh@ing.unlpam.edu.ar}

Resumen Este proyecto de investigación se enfoca en la resolución de problemas de optimización a gran escala utilizando nuevas técnicas metaheurísticas, así como también su hibridación con las ya existentes.

Una de las líneas de investigación analiza el efecto de reemplazar el método para crear nuevas soluciones en el algoritmo *artificial bee colony* por operadores de recombinación.

Otra de las líneas de investigación se enfoca en la resolución del problema *flexible job shop scheduling* (NP-hard), presente en ambientes fabriles, porque tiene que asignar cada operación a la máquina apropiada además de secuenciar las operaciones en las máquinas. Debido a esta complejidad, las metaheurísticas se convirtieron en la mejor opción para resolver en la práctica este problema.

Una tercera línea de investigación apunta a la resolución del problema de diseño de redes de distribución de agua, mediante el uso de metaheurísticas como *Simulated Annealing* y *Cuckoo Search*.

Por último, una línea de investigación se orienta a la utilización de la metaheurística basada en la migración de las aves en el problema de ruteo vehicular con capacidad, el cual es reconocido por su incidencia en el mundo de los negocios y por la dificultad para resolverlo.

Palabras claves: Metaheurísticas, optimización, operadores

Contexto

Estas líneas de investigación se desarrollan en el marco de un proyecto de investigación, llevado a cabo en el Laboratorio de Investigación de Sistemas Inteligentes (LISI) de la Facultad de Ingeniería de la Universidad Nacional de La Pampa, acreditado por dicha facultad y dirigido por la Dra. Minetti. Cabe destacar que desde hace varios años, los integrantes de estos proyectos mantienen una importante vinculación con investigadores de la Universidad Nacional de San Luis (Argentina) y de la Universidad de Málaga (España), con quienes se realizan publicaciones conjuntas.

Introducción

La mayoría de los problemas de optimización del mundo real abordan una gran cantidad de variables de decisión, y se los conoce como problemas de Optimización Global a Gran Escala (LSGO). Muchas aplicaciones de la ciencia y la ingeniería se formulan como problemas LSGO: diseño de sistemas electrónicos a gran escala, problemas de *scheduling* con gran cantidad de recursos u operaciones, diseño de redes de distribución de agua, ruteo vehicular a gran escala, reconocimiento de genes en bioinformática, cinética química de problemas inversos, entre otros. Por un lado, la optimización de este tipo de problemas consume mucho tiempo. Por otra parte, los algoritmos metaheurísticos son ampliamente reconocidos como enfoques eficientes para resolver problemas de optimización de grandes dimensiones. En este sentido,

las líneas de investigación de este proyecto se encargan de proponer, adaptar y analizar distintas metaheurísticas con el propósito de resolver eficaz y eficientemente diferentes problemas LSGO.

Por un lado, se estudia y analiza el reemplazo del método para crear nuevas soluciones en el algoritmo de *artificial bee colony* mediante operadores de recombinación. Las razones de este reemplazo están dadas por el hecho que el método original es similar al proceso de recombinación utilizado en los algoritmos evolutivos. Para ello, realizamos una investigación sistemática del efecto de utilizar seis operadores de recombinación diferentes, para soluciones codificadas en los reales, en el paso realizado por la abeja empleada.

Por otra parte, el problema *flexible job shop scheduling* (FJSSP) es un problema difícil y complejo en entornos de fabricación, NP-hard, porque tiene que asignar cada operación a la máquina adecuada además de secuenciar las operaciones en tales máquinas. Debido a esta complejidad, las metaheurísticas se convirtieron en la mejor opción para resolver en la práctica este tipo de problema. Por lo tanto, se trabajó en la realización de un compendio cubriendo un amplio espectro algorítmico de diferentes técnicas para resolver el JSSP, analizando la calidad de las soluciones encontradas por cada una de estas técnicas, además de las ventajas y desventajas de las mismas en cuanto a tiempos de desarrollo, adecuación al problema, complejidad en la resolución del problema, entre otras características.

Además se investiga la optimización del diseño de redes de distribución de agua, que es un campo de investigación muy activo desde hace algunas décadas. Estas redes están compuestas por reservorios y tuberías que tratan de brindarles a los usuarios un flujo constante de agua con una determinada presión. El problema de optimizar estas redes consiste en encontrar el diámetro óptimo de cada tubería seleccionándola de un conjunto limitado de caños disponibles comercialmente, con el objetivo de reducir el costo. En esta línea de investigación se están utilizando técnicas metaheurísticas basadas en trayectoria y en población. En particular se están adaptando y probando diferentes algoritmos como *Simulated Annealing* (SA), *Cuckoo Search* (CS), entre otros.

Por último, en otra línea de investigación se trabaja sobre el algoritmo propuesto por Duman [1] basado en la migración de las aves (*Migrating Birds Optimization Algorithm*, MBO). Este algoritmo ha demostrado ser eficiente al resolver problemas de transporte marítimo [2], secuenciación de tareas [3], entre otros. Motivo por el cual se diseñó e implementó una versión de MBO para resolver el problema de ruteo vehicular (CVRP). El objetivo es evaluar la calidad de las soluciones encontradas por esta propuesta algorítmica y analizar su esfuerzo computacional. Esta nueva versión de MBO es testada usando un amplio y diverso conjunto de casos de prueba.

Desarrollo

En esta sección se describen las líneas de investigación mencionadas en la introducción.

Mejora del algoritmo *artificial bee colony*

La inteligencia *swarm* es el estudio de sistemas computacionales inspirados por la cooperación de un gran número de agentes homogéneos en el medio ambiente. Una colonia de hormigas o de abejas, una bandada de pájaros, o un sistema inmune son ejemplos típicos de sistemas de enjambres. Tereshko y Loengarov [4] consideran una colmena de abejas como un sistema dinámico, donde recopilar información de un entorno y ajustar su comportamiento de acuerdo con ésta produce una toma de decisiones inteligente al mejorar el nivel de comunicación entre los individuos. Teniendo en cuenta las ideas de Tereshko y Loengarov, Karaboga et al. [5] proponen un algoritmo de optimización basado en el comportamiento inteligente del enjambre de abejas (*artificial bee colony*, ABC).

En la literatura se presentan alternativas que modifican el método original para generar una fuente de comida candidata (o solución candidata) [6], [7], [8], [9], pero estas alternativas son solo pequeñas variaciones del esquema tradicional. El mecanismo utilizado por ABC para producir una nueva solución candidata es muy similar al procedimiento llevado a cabo por los operadores de recombinación en la literatura

de los algoritmos evolutivos. En este sentido, se consideró la aplicación de otros mecanismos para generar nuevas posiciones de fuentes al utilizar los operadores de recombinación, cuyo efecto sobre el desempeño de ABC aún no se ha estudiado y su impacto podría ser más significativo que el enfoque tradicional. En consecuencia, se analizó el efecto del uso de seis operadores de recombinación diferentes para crear nuevas posiciones de alimentos en el paso de la abeja empleada del algoritmo ABC. Dado que ABC trabaja con soluciones codificadas en los reales, se eligieron los siguientes operadores evolutivos: aritmético, aritmético-Max-Min, lineal, binomial, de un punto y multipunto. El desempeño de estas variantes algorítmicas se evaluó al usar el conjunto de prueba IEEE CEC'2008 [10], especialmente diseñado con problemas de minimización de gran escala (es decir con dimensiones de 100, 500 y 1000 variables).

Optimización del diseño de redes de distribución de agua

Se requiere un método de solución eficaz que sea confiable y fácil de usar para la optimización de las redes de distribución de agua, que proporcionan un servicio esencial en todas las comunidades. La optimización no solo aborda los costos de capital y operativos junto con el rendimiento y la confiabilidad hidráulica, sino también la gestión competente de la energía. En consecuencia, las metaheurísticas brindan, una vez más, una alternativa de solución eficiente. Por este motivo, en esta línea de investigación se analizan y diseñan dos variantes metaheurísticas que resuelven este problema, SA e ILS.

Dada las características propias de este tipo de redes, una vez que SA o ILS arman una determinada solución, la prueban y evalúan por medio del simulador EPANET 2.0 [11]. Este simulador, además, resuelve todas las ecuaciones hidráulicas de forma externa. Los problemas que se utilizan para efectuar los experimentos son de periodo simple y multi periodo, en donde el patrón de demanda varía con el tiempo. Si bien, aún, no se cuenta con resultados fehacientes, por estar en la fase inicial, se puede observar

que los algoritmos desarrollados logran un buen desempeño al reducir en forma considerable los costos de las redes.

Compendio de metaheurísticas eficientes para resolver el FJSSP

Generalmente, el FJSSP es un entorno de producción más realista y, por consiguiente, en la práctica es más aplicable que el JSSP. Sin embargo, el FJSSP es más complejo [12]) que el JSSP debido a su decisión adicional de asignar cada operación a la máquina apropiada (enrutamiento), además de, secuenciar las operaciones en estas máquinas (planificación o *scheduling*). Debido a la complejidad mencionada de FJSSP, se sugiere la adopción de métodos heurísticos porque producen *schedules* de muy buena calidad en un tiempo razonable, en lugar de buscar una solución óptima, posible sólo para casos pequeños. En los últimos años, la adopción de metaheurísticas [13] ha dado lugar a mejores resultados que los algoritmos exactos o los heurísticos *greedy* [14], [15], [16].

Por todo lo anterior, se adaptaron y analizaron diferentes metaheurísticas para resolver el FJSSP con el objetivo de ofrecer una referencia de base para la comparación, utilizando diferentes metaheurísticas para cubrir un amplio espectro algorítmico, desde los métodos de trayectoria hasta los métodos basados en la población. De esta forma, obtenemos un valioso compendio de enfoques para FJSSP al especificar sus cualidades desde dos puntos de vista: primero, adaptando las metaheurísticas para resolver el problema y, en segundo lugar, encontrando *schedules* de buena calidad con un bajo esfuerzo computacional. En consecuencia, la metodología seguida en este estudio empírico consiste en adaptar cinco metaheurísticas para resolver el FJSSP: SA, *Iterated Local Search* (ILS), CS, *Genetic Algorithms* (GA) e *Imperative Competitive Algorithm* (ICA). Se evaluaron y compararon cada una de ellas en un conocido conjunto de instancias de prueba del problema.

Aplicación de MBO al problema de ruteo vehicular

El problema de ruteo vehicular con capacidad (CVRP) es ampliamente conocido por su complejidad y por su vasta aplicación en la logística empresarial. Muchas han sido las técnicas algorítmicas usadas para resolverlo y MBO es una que aún no había sido considerada para tal fin. MBO [1] toma como referencia el comportamiento de las aves durante sus largos viajes migratorios y lo reformula en un algoritmo de optimización. Aquí se trata a las soluciones como aves migratorias que emprenden el vuelo, por ende el primer paso consiste en generar de forma aleatoria las aves que compondrán la formación y ubicarlas de manera hipotética en forma de "V". A partir de estas soluciones iniciales, el algoritmo comenzará la exploración y explotación del espacio de búsqueda, simulando la rotación de las aves en la formación de vuelo. El trabajo de investigación consistió en adaptar este algoritmo al CVRP, al modificar la representación de la solución y los operadores de movimiento; los cuales diferían de los definidos originalmente para MBO. Esto significó modificaciones en la inicialización de la población, en la definición de la estructura de vecindario y en su manejo.

Resultados obtenidos

A continuación se detallan los resultados obtenidos por cada una de las líneas de investigación abordadas, excepto en la referida a la optimización de redes de distribución de agua por encontrarse en una fase inicial.

Los resultados obtenidos, al analizar los seis operadores de recombinación diferentes para crear nuevas fuentes de alimentos en el algoritmo ABC, sugieren el uso de la recombinación aritmética como una alternativa interesante al método original [17]. Dado que este operador presenta un rendimiento muy bueno en términos del rango de calidad promedio y el esfuerzo computacional para todas las funciones y dimensiones consideradas. Además, el desempeño de los otros operadores que ajustan las variables en las soluciones (operadores de recombinación lineal y aritmético Max-Min) es similar al comportamiento del algoritmo origi-

nal. En cambio, los algoritmos ABC que aplican operadores de recombinación binomial, de un solo punto y multipunto, y solo copian partes de otras soluciones para crear uno nuevo, no parecen ser buenas alternativas para el algoritmo ABC, debido a la baja calidad de sus soluciones para cualquier función y dimensión.

En cuanto al compendio de diferentes metaheurísticas para resolver el FJSSP, se incluyeron dos metaheurísticas basadas en trayectoria (SA e ILS) y tres basadas en población (CS, GA e ICA) para dar un amplio espectro de posibles soluciones al problema de programación mencionado. Los resultados indican que las metaheurísticas basadas en trayectoria fueron precisas y eficientes para solucionar el FJSSP [18]. En el estudio llevado a cabo en este trabajo, SA es el algoritmo de mejor rendimiento para resolver el FJSSP. Además, cuando SA se contrasta con algoritmos en la literatura, también se convierte en el mejor enfoque. Como consecuencia, SA brinda soluciones precisas a este problema de NP-hard de una manera eficiente y competitiva.

El desempeño de MBO fue puesto a prueba con numerosas instancias de CVRP de variada complejidad, obteniendo resultado de buena calidad en tiempos razonables. Esto motiva a continuar trabajando en la mejora del método de exploración de MBO, con el fin de alcanzar la solución óptima en la mayor cantidad de instancias posibles. Es decir que, dado a los bajos tiempos de cómputo de la metaheurística se dispone de un margen de acción que permite llevar a cabo esta tarea y mantener valores de tiempo competitivos.

Formación de recursos humanos

Cada año se incorporan al proyecto alumnos avanzados en la carrera Ingeniería en Sistemas, quienes trabajan en temas relacionados a la resolución de problemas de optimización usando técnicas inteligentes, con el objeto de guiarlos en el desarrollo de sus tesis de grado y, también, de formar futuros investigadores científicos. Por otra parte, los docentes-investigadores que integran el proyecto realizan diversos cursos de posgrado relacionados con la temática del proyecto, con

el objetivo de sumar los créditos necesarios para cursar carreras de posgrado.

REFERENCES

- [1] E. Duman, M. Uysal, and A. F. Alkaya, "Migrating birds optimization: A new metaheuristic approach and its performance on quadratic assignment problem," *Information Sciences*, vol. 217, pp. 65 – 77, 2012.
- [2] E. Lalla-Ruiz, C. Expósito-Izquierdo, J. de Armas, B. Melián-Batista, and J. M. Moreno-Vega, "Migrating birds optimization for the seaside problems at maritime container terminals," *Journal of Applied Mathematics*, vol. 2015, Special Issue, p. 12 pages, 2015. [Online]. Available: <https://doi.org/10.1155/2015/781907>
- [3] "The migrating birds optimization metaheuristic for the permutation flow shop with sequence dependent setup times," *IFAC-PapersOnLine*, vol. 49, no. 12, pp. 408 – 413, 2016.
- [4] V. Tereshko and A. Loengarov, "Collective decision-making in honey bee foraging dynamics," *Computing and Information Systems Journal*, ISSN 1352-9404, 2005.
- [5] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (abc) algorithm," *J. of Global Optimization*, vol. 39, no. 3, pp. 459–471, Nov. 2007.
- [6] W. Gao, S. Liu, and L. Huang, "A global best artificial bee colony algorithm for global optimization," *Journal of Computational and Applied Mathematics*, vol. 236, no. 11, pp. 2741 – 2753, 2012.
- [7] K. Diwold, A. Aderhold, A. Scheidler, and M. Middendorf, "Performance evaluation of artificial bee colony optimization and new selection schemes," *Memetic Computing*, vol. 3, no. 3, p. 149, 2011.
- [8] B. Akay and D. Karaboga, "A modified artificial bee colony algorithm for real-parameter optimization," *Information Science*, vol. 192, pp. 120–142, 2012.
- [9] A. Banharsakun, T. Achalakul, and B. Sirinaovakul, "The best-so-far selection in artificial bee colony algorithm," *Applied Soft Computing*, vol. 11, no. 2, pp. 2888 – 2901, 2011, the Impact of Soft Computing for the Progress of Artificial Intelligence.
- [10] K. Tang, X. Yao, P. N. Suganthan, C. MacNish, Y. P. Chen, C. M. Chen, , and Z. Yang, "Benchmark functions for the CEC'2008 special session and competition on large scale global optimization," Nature Inspired Computation and Applications Laboratory, USTC, China, Technical Report, 2007.
- [11] L. Rossman, *Users Manual*, Environmental Protection Agency: Washington, DC, USA, 2000. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2405896316309132>
- [12] M. R. Garey, D. S. Johnson, and R. Sethi, "The complexity of flowshop and jobshop scheduling," *Math. Oper. Res.*, vol. 1, no. 2, pp. 117–129, May 1976.
- [13] E. Talbi, *Metaheuristics: from Design to Implementation*. Wiley, 2009.
- [14] G. Vilcot and J.-C. Billaut, "A tabu search and a genetic algorithm for solving a bicriteria general job shop scheduling problem," *European Journal of Operational Research*, vol. 190, no. 2, pp. 398 – 411, 2008.
- [15] G. Zhang, X. Shao, P. Li, and L. Gao, "An effective hybrid particle swarm optimization algorithm for multi-objective flexible job-shop scheduling problem," *Computers & Industrial Engineering*, vol. 56, no. 4, pp. 1309 – 1318, 2009.
- [16] G. Zhang, L. Gao, and Y. Shi, "An effective genetic algorithm for the flexible job-shop scheduling problem," *Expert Syst. Appl.*, vol. 38, pp. 3563–3573, 2011.
- [17] G. Minetti and C. Salto, "Improving artificial bee colony algorithm with evolutionary operators," in *Anales del XXIII Congreso Argentino de Ciencias de la Computación*, R. de Universidades con Carreras en Informática (RedUNCI), Ed., 2017, pp. 93–102.
- [18] C. Bermudez, G. Minetti, and C. Salto, "A metaheuristic compendium for scheduling problems," in *Anales del XXIII Congreso Argentino de Ciencias de la Computación*, R. de Universidades con Carreras en Informática (RedUNCI), Ed., 2017, pp. 2–12.

Desarrollo de Sistemas de Scheduling para la Nube

Daniel Díaz, Francisco Ibañez, Sandra Oviedo, Nancy Alonso, Juan Manuel Cuneo

LISI- Instituto de Informática – Dpto. de Informática
FCEFYN - Universidad Nacional de San Juan
CUIM – Av. Ignacio de la Roza 590 (O), Rivadavia – J5402DCS San Juan
{ fibanez, ddiaz, soviedo, nalonso, jmcuneo}@iinfo.unsj.edu.ar}

Resumen

Los sistemas de scheduling son el corazón de todo sistema de planificación de la producción. Las nuevas tecnologías traen consigo amenazas y oportunidades para la industria, en este contexto, los sistemas de scheduling no escapan a esta realidad. El objetivo de este proyecto es investigar sobre métodos y técnicas en pos de integrar las tecnologías de cloud computing, web de tiempo real, software product lines y de los solucionadores para optimización en el desarrollo de sistemas de scheduling de producción.

Palabras clave:

Scheduling de producción, cloud computing, web de tiempo real

Contexto

Este trabajo se desarrolla en el marco del proyecto denominado “Desarrollo de sistemas de scheduling producción utilizando computación en la nube, web de tiempo real, y línea de productos de software”. La unidad ejecutora es el Laboratorio de Informática Aplicada a la Innovación del Instituto de informática de la Facultad de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de San Juan.

Introducción

Entre las últimas tecnologías que se han instalado en el mercado, la computación en la nube (cloud computing) es una tecnología que está permitiendo que las pequeñas y medianas empresas utilicen los software que en la década pasada estaba solo permitidos, por sus costos de licenciamiento, a las grandes empresas. Las herramientas informáticas que permiten solucionar problemas de scheduling se denominan “optimization solvers”, son plataformas que permiten modelar, diseñar e implementar soluciones para problemas de optimización. Los problemas de scheduling son un caso particular de los problemas de optimización. Muchas de estas herramientas están desplegadas en la nube como servicios. Web de tiempo real se refiere a la capacidad que tiene una aplicación web de comunicar diferentes usuarios o dispositivos de manera tal que la información está disponible tan pronto es publicada por su autor, es decir no requiere que cada participante observe periódicamente la fuente para ver si existen información nueva o modificada.

Software product lines es una tecnología que promueve la reutilización del software para el desarrollo de familias

de productos, recientemente estas técnicas se están aplicando a la construcción de aplicaciones que se ejecutan en el modelo de cloud computing, dicha tecnologías se conocen como cloud-based software product lines.

El resultado de la integración de estas tecnologías debe permitir disponer de sistemas scheduling que se ejecuten en la nube, hagan uso de la web de tiempo real, y aprovechen todos los beneficios que una línea de producto de software brinda en cuanto a customización y reducción de costos de implementación y mantenimiento.

Sistemas de scheduling en dominios industriales

Los problemas de Scheduling aparecen en diferentes dominios, este proyecto trata con problemas de scheduling que surgen en el dominio industrial y que en la industria que se denominan *production scheduling* o *manufacturing scheduling*.

La definición más clásica de la palabra scheduling dice : “Scheduling es el problema de asignar recursos limitados a tareas en el tiempo con el objeto de optimizar uno o más objetivos” [1]. En la industria, un sistema de scheduling es el corazón del sistema de planificación y control de la producción. Normalmente el sistema de scheduling interactúa con otros sistemas de fabricación, tales como ERP (Enterprise Resource Planning) y MES (Manufacturing Execution System).

Existe una literatura muy extensa sobre scheduling y se ha escrito mucho sobre modelos y algoritmos que resuelven problemas de scheduling. Sin embargo, poco son los artículos que tratan de cómo traer estos modelos y algoritmos a implementaciones reales. Esto se conoce como el “hueco o gap” entre la teoría y la práctica de scheduling [2]. Para cerrar este hueco entre la teoría y la práctica los

modelos y algoritmos de scheduling se deben implementar en una pieza de software que satisfaga las necesidades que tiene una empresa en el área de scheduling. Esto implica llevar a cabo un proceso de desarrollo de software para obtener un producto final, es decir un sistema de scheduling y que ayude a la toma de decisión. Según Yen y Pinedo [3] un sistema de scheduling se compone tres módulos: (1) módulos de base datos y base de conocimiento, (2) módulos del motor de scheduling y (3) módulo de interface de usuario. La parte esencial de todo sistema de scheduling es el desarrollo del motor de scheduling, es donde se vinculan las necesidades que tiene la empresa con los modelos y algoritmos de scheduling.

Debido a los avances tecnológicos y evolución del mercado nuevas necesidades por parte de la empresa han surgido. McKay et al. [4] describen un conjunto de necesidades o requisitos que debe reunir un sistema de scheduling, algunos de los cuales se refieren a los algoritmos de secuenciación y al uso del sistema. Con respecto a los algoritmos de secuenciación, éstos deben ser flexibles y configurables. Deben ser capaces de manejar diversos criterios (scheduling multicriterio). A estos requisitos se les debe agregar otros tales como facilitar el mantenimiento y la reducción de costos de desarrollo.

Por último , [5] intenta integrar esta definiciones definiendo a un manufacturing scheduling como el proceso de toma de decisión que consiste de asignar un conjunto de operaciones/tareas requeridas para manufactura un conjunto de productos con los recursos existentes en una planta, como así también los tiempos necesario para iniciar estas operaciones o tareas. Un Schedule o plan se define como un conjunto específico de asignaciones de

operaciones o tareas a los recursos sobre una escala temporal.

Evolución de los sistemas de scheduling

La investigación en el diseño y desarrollo de sistemas de scheduling se puede clasificar en 5 generaciones. Inicialmente, esta clase de problemas fue abordada por el área de la ingeniería industrial, y posteriormente por el dominio de la investigación operativa. En los comienzos del siglo pasado, Henry L. Gantt, fue uno de los primeros en proponer un sistema de scheduling: daily balance “el equilibrio diario” que consiste en un método de programar y de registrar el trabajo. En 1960 se desarrolló Planalog Control Board, una herramienta que constaba de una pizarra y entre otras cosas podía forzar restricciones de precedencia. Los primeros sistemas de scheduling computarizados nacieron con el MRP (Material Requirement Planning). En la segunda generación surgen los ISS (Intelligent Scheduling Systems), tales como ISIS y OPIS de [6]. Esta fue una etapa donde el problema de scheduling se comenzó a abordar con técnicas de inteligencia artificial tales como algoritmos genéticos, simulated annealing, redes neuronales y constraint satisfaction techniques. En esta generación aparecen los sistemas de scheduling que se construyen utilizando técnicas de KBS (Knowledge Based Systems). La tercera generación se inició con el surgimiento del paradigma orientado, dando origen a sistemas de scheduling desarrollados con esta tecnología. Esto permitió el desarrollo de frameworks especializados en scheduling, los cuales encapsulan técnicas de inteligencia artificial [7]. La cuarta generación de sistemas de scheduling es la que se integra en lo que se conoce

como Advanced Planning and Scheduling Systems [8]. La quinta generación está surgiendo actualmente, está basada sobre la reutilización sistemática del software, la computación en la nube y la web de tiempo real.

Computación en la nube.

Según [9] Cloud Computing es un modelo para habilitar acceso conveniente por demanda a un conjunto compartido de recursos computacionales configurables, por ejemplo, redes, servidores, almacenamiento, aplicaciones y servicios, que pueden ser rápidamente provisionados y liberados con un esfuerzo mínimo de administración o de interacción con el proveedor de servicios. La característica principal de este modelo es que el software se despliega como servicios que están alojados en la nube, aquí los usuarios pagan por lo que usan, es decir son servicios elásticos. Esta característica permite una escalabilidad flexible de un sistema basado en el modelo de cloud computing, es decir un sistema puede atender a un solo usuario en un momento y a miles de usuarios en otro momento, los costes de la atención varían en función de la cantidad de usuarios que hay para un momento dado. Normalmente en este modelo atender pocos usuarios resulta mucho más económico que comprar las licencias de un software.

Programación web en tiempo real

La web en tiempo real es la web en la que los datos se entregan a sus destinatarios (sean humanos o máquinas) en tiempo real o casi real, tan pronto como estén disponibles. Su alcance se extiende desde las finanzas y herramientas médicas, a los servicios de redes sociales y medios de comunicación. Los principales requisitos y capacidades de aplicaciones en tiempo real son

interactividad para presentar los datos en tiempo real a los clientes, capacidad para mantener una conexión permanente con el servidor, protocolo de transmisión de datos más rápida, capacidad para integrar la lógica del cliente para reducir la carga del servidor, capacidad para enviar datos de múltiples fuentes, incluyendo los cambios de otros usuarios [10].

Meteor es un framework de JavaScript que tiene como objetivo automatizar y simplificar el desarrollo de aplicaciones web que actúan en tiempo real. Usa la comunicación en tiempo real mediante un protocolo llamado Distributed Data Protocol (DDP), que es admitido en navegadores nuevos que usan WebSockets o en navegadores anteriores que usan Asynchronous JavaScript + XML (Ajax) de sondeo largo. En ambos casos, la comunicación navegador-servidor es transparente [11].

Líneas de producto de software

La Línea de Productos de Software (LPS) se refiere a técnicas de ingeniería para crear sistemas de software que comparten un conjunto común y gestionado de características que satisfacen las necesidades específicas de un segmento de mercado y son desarrollados a partir de un conjunto compartido de activos de software, usando un medio común de producción [12-14]. Los beneficios que se tienen al hacer uso de la LPS es la entrega de productos de software de una manera más rápida, económica y con una mejor calidad.

El Desarrollo de Software Orientado a Características (DSOC) es un paradigma para la construcción, la personalización, y la síntesis de Sistemas de Software. Una característica es un aspecto de un sistema de software distintivo o visible para el usuario. La idea básica de DSOC es descomponer los sistemas de software en características con el fin de proporcionar

opciones de configuración y facilitar la generación de sistemas de software basado en una selección de características [15]. De esta manera, una línea de productos de software se refiere a un conjunto de sistemas de software que pueden ser obtenidos a partir de un determinado conjunto de características. Un determinado código (o parte) es mapeado en una característica con el objeto de construir un sistema de software a medida con la única acción de seleccionar las características que se desean de este software. FeatureIDE es un marco de trabajo para el Desarrollo de Software Orientado a Características (DSOC) basado en Eclipse. Según [16], FeatureIDE se enfoca en todo el proceso de desarrollo e incorpora herramientas para la implementación de Líneas de Producto de Software en un Entorno de Desarrollo Integrado. FeatureIDE soporta la implementación de varias técnicas como la programación orientada a características, la programación orientada a aspectos y preprocesadores.

Resultados y Objetivos

El objetivo es investigar sobre métodos y técnicas en pos de integrar las más recientes tecnologías de cloud computing, web de tiempo real, software product lines y de los solucionadores para optimización en el desarrollo de sistemas de scheduling de producción.

Estas nuevas tecnologías presentan nuevos desafíos y oportunidades para los sistemas de toma de decisión y optimización en particular para los sistemas de planificación y scheduling de la producción. La toma de decisión y optimización es una línea de investigación en la que el grupo de investigación cuenta con mucha experiencia. Por otro lado, nuestro grupo de investigación en los últimos años

también ha estado trabajando en otros proyectos relacionados con estas tecnologías desde que las mismas se consideraban emergentes. Creemos que las tecnologías ya están lo suficientemente maduras como para ser incorporadas a los sistemas de toma de decisión y optimización.

Formación de Recursos Humanos

El equipo de trabajo que lleva adelante este proyecto se compone de

- 5 docentes investigadores,
- 3 tesistas de grado en período de finalización.
- 2 tesistas de grado en período iniciación.
- 2 tesistas de posgrado (maestría) iniciando sus trabajos.

Referencias

- [1] K. R. Barker, *Elements of sequencing and scheduling*. New York: John Wiley and Sons, 1974.
- [2] B. L. MacCarthy and J. Liu, "Addressing the gap in scheduling research: A review of optimization and heuristic methods in production scheduling," *International Journal of Production Research*, vol. 31, pp. 59-79, 1993.
- [3] B. P.-C. Yen and M. Pinedo, "On the design and development of scheduling systems," in *Fourth International Conference on Computer Integrated Manufacturing and Automation Technology*, 1994, pp. 197 - 204.
- [4] K. McKay, M. L. Pinedo, and S. Webster, "Practice-Focused Research issues for scheduling systems," *Production and Operations Management*, vol. 11, pp. 249-258, 2002.
- [5] J. M. Framinan, R. Leisten, and R. R. García, "Manufacturing scheduling systems," *An integrated view on Models, Methods and Tools*, pp. 51-63, 2014.
- [6] F. S. Smith, *OPIS: A Methodology and architecture for reactive scheduling*: Morgan Kaufmann., 1994.
- [7] C. Le Pape, "Implementation of Resource Constraints in ILOG SCHEDULE: A Library for the Development of Constraint-Based Scheduling Systems," *Intelligent Systems Engineering* 3, vol. 3, pp. 55-66., 1994.
- [8] M. L. Entrup, *Advanced Planning and Scheduling Systems: Advanced Planning in Fresh Food Industries: Integrating Shelf Life into Production Planning*. Physica-Verlag HD, 2005.
- [9] N. Antonopoulos and L. Gillam, *Cloud computing*: Springer, 2010.
- [10] D. Berube, "Desarrollo de sitios web en tiempo real " *IBM-Developers Works*, vol. <http://www.ibm.com/developerworks/ssa/library/wa-meteor/wa-meteor-pdf.pdf>, 2013.
- [11] F. Vogesteller, "Building Single-page Web Apps with Meteor," *Packt Publishing*, 2015.
- [12] J. Greenfield and K. Short, "Software Factories: Assembling Applications with Patterns, Models, Frameworks, and Tools," *Wiley*, 2004.
- [13] K. Pohl, G. Böckle, and F. J. von der Linden, "Software Product Line Engineering: Foundations, Principles and Techniques," *Springer* 2005.
- [14] D. M. Weiss and C. T. R. Lai, "Software Product-Line Engineering - A Family-Based Software Development Process.," *Addison-Wesley. Massachusetts*, 1999.
- [15] K. Czarnecki, "Overview of Generative Software Development," *Unconventional Programming Paradigms*, 2005.
- [16] C. Kastner, T. Thum, G. Saake, and J. Feigenspan, "FeatureIDE: A tool framework for feature-oriented software development " *IEEE 31st ICSE* 2009.

Modelo de Inteligencia Artificial aplicados al Trading Algorítmico

Hernán Merlino, Eduardo Diez, Hernán Amatriain, Nicolás Pérez, Santiago Bianco, Sebastian Martins

Laboratorio de Investigación y Desarrollo en Ingeniería de Explotación de Información
Grupo de Ingeniería de Explotación de Información y Grupo Investigación en Sistemas de Información

Departamento de Desarrollo Productivo y Tecnológico, Universidad Nacional de Lanús
29 de Septiembre 3901 (1826) Remedios de Escalada, Lanús, Argentina. Tel +54 11 5533 5600 Ext. 5194
hmerlino@gmail.com, hamatriain@gmail.com

RESUMEN

El constante aumento en el uso de modelos computacionales en la industria financiera específicamente en el trading de valores [Boehmer, Fong & We, 2015; Chaboud, Hjalmsrsson, 2015] ha traído como consecuencia errores en los modelos dado a la mala codificación de los programas involucrados.

El objetivo general del proyecto es el desarrollo de un modelo de inteligencia artificial que tome decisiones de compra y venta de valores bursátiles en forma automática (trading algorítmico) y pueda presentar una performance similar a un operador humano. De forma adicional, para la construcción del modelo propuesto se realizará una experimentación controlada, haciendo uso de los patrones de error más comunes, incorporando a la propuesta un modelo de auditoria para la evaluación de los algoritmos utilizados en trading algorítmico.

Palabras clave: Inteligencia Artificial, Trading Algorítmico, Valores Bursátiles, Formación de Recursos Humanos.

CONTEXTO

El Proyecto articula líneas de investigación del Grupo de Investigación en Sistemas de Información y el Laboratorio de Investigación y Desarrollo en Arquitecturas Complejas de la Licenciatura en Sistemas del Departamento de Desarrollo Productivo y Tecnológico de la Universidad Nacional de Lanús. Dentro de la mencionada Universidad, el Proyecto está inscripto con el número 80020170200007LA.

FUNDAMENTACION

En la actualidad el 75% de las transacciones bursátiles que se hacen en el mundo son ejecutadas por sistemas computacionales, comúnmente llamados robots; es aquí donde se ha detectado una vacancia de conocimiento que amerita la dedicación de recursos y esfuerzos al mismo, pues estos robots son artefactos software codificados por un programador, esto hace que exista la probabilidad de errores en el código que pueden producir colapsos en el sistema financiero. Esto ya ha pasado en el pasado reciente, flash crash 2010, pero lo más importante que se incrementarán las probabilidades, pues cada vez más se utiliza estos modelos computacionales.

Se han hecho estudios sobre el análisis de sentimiento para la implementación de modelos de este tipo [Kleinnijenhuis, Schultz y Oegema, 2013], permitiendo incluir a la opinión de internautas en los modelos de toma de decisiones sobre compra y venta de valores. Por otra se ha trabajado en el descubrimiento de reglas aplicando modelos evolutivos [Hua, Liu, Zhanga, 2015] que interactúan con los modelos basados en sentimientos.

También se ha encontrado evidencia sobre trabajos en modelos de evaluación de performance de trading algorítmico [Cooper, Ong & Van Vliet, 2015], los cuales permiten una vez estos se encuentren en producción medir su rendimiento; a esto se le agrega los modelos de evaluación de la calidad de los mercados y el trading algorítmico [Scholtus, Dijk, Frijns, 2014], estos últimos hacen especial referencia a la volatilidad que inyectan al mercado los modelos de trading algorítmico. También se ha trabajado en modelos para la detección de precios [Brogaard, Hendershott, Riordan 2014; Viljoen, Westerholm & Zheng

2016; Brogaard, Hendershott, Riordan, 2016], haciendo un intento de detectar la volatilidad dada por los mercados.

Otro conjunto de investigadores han trabajado en técnicas para la selección de portfolio [Berutich, López & Luna, 2016] mediante modelos computacionales, estos intentan obtener una cartera de inversiones optimas sin el sesgo cognitivo que suelen aplicar los seres humanos a estas.

Un aporte sustancial al problema en cuestión lo ha hecho Jakob Arnoldi [Arnoldi, 2015] demostrando como el trading algorítmico modifica hoy en día el precio de la acciones, bonos o mercado en el que se encuentre operando este tipo de modelos; por otra parte, Eric Allen [Allen, 2002] ha descripto las principales características y patrones en los errores en la codificación de software.

Continuando con la línea de investigación en el área del trading algorítmico, se pueden mencionar aportes por parte de [Hasbrouck & Gideon, 2013; Evans, Pappas & Xhafa, 2013; Tarun et.al., 2013; Arnoldi, 2015; Nuij et.al., 2015] donde hacen mención a los cambios producidos en los precios de los valores por el uso de modelos de trading algorítmico.

Es de destacar que al momento de llevar adelante el proceso de revisión sistemática para este proyecto no se ha podido detectar trabajos que integren el trading algorítmico y la detección de errores en el código de los mismos. En este contexto, se define como objetivo general del presente proyecto es el desarrollo de un modelo de inteligencia artificial que tome decisiones de compra y venta de valores bursátiles en forma automática (trading algorítmico) y pueda presentar una performance similar a un operador humano.

La relevancia de la investigación está dada en tres sentidos en el área científica al momento de realizar la investigación documental no se ha encontrado una aproximación de este tipo al problema, si similares como se detalla en los siguientes apartados. Social, las próximas crisis económicas no se producirán por guerras o desastres naturales, sino por los errores en la forma de hacer transacciones automática, y la sociedad en su conjunto pedirá explicaciones sobre porque han perdido sus ahorros, es ahí

donde la universidad deberá dar explicaciones y en el mejor de los casos, advertir antes que esto suceda. El tercer pilar es en el terreno educativo, cuanto más se sepa sobre este tema, se lo podrá controlar mejor.

OBJETIVOS Y METODOLOGÍA DE DESARROLLO

Objetivo general o marco de referencia

El problema al que se enfrenta la sociedad en su conjunto es, como se dijo en el apartado anterior, el 75% de las transacciones bursátiles que se hacen en el mundo son ejecutadas por sistemas computacionales, en consecuencia es muy alto el porcentaje de errores de codificación al que están expuestos los sistemas financieros.

Por esto, el objetivo principal es el del desarrollo de un modelo de inteligencia artificial para la toma de decisiones en la compra y venta de acciones y otros valores en la bolsa de comercio, este nos permitirá generar las simulaciones necesarias para reconocer los patrones de error y poder hacer un aporte sustancial al área de incumbencia del mismo.

Estrictamente, los aportes que se espera de esta investigación son: (a) modelo de simulación de algoritmos basados en inteligencia artificial para compra y venta de acciones y otros valores en la bolsa de comercio; (b) Detección de patrones de errores en este modelo de algoritmo y (c) presentar un modelo de auditoria que permite reducir el impacto ante la presencia de estos en los sistemas de trading algorítmico.

Objetivos Específicos:

Objetivo Especifico 1: Desarrollar un Proceso de Construcción de Algoritmos Basados en Inteligencia Artificial para la toma de decisiones en la compra y venta de Acciones en la bolsa de comercio.

Objetivo Especifico 2: Reconocimiento de los patrones de errores más comúnmente realizadas

en el desarrollo de modelos de trading algorítmico.

Objetivo Especifico 3: Desarrollar un Proceso de Auditoria para la evaluación de los algoritmos utilizados en trading algorítmico.

Metodología:

Para construir el conocimiento asociado al presente proyecto de investigación, se seguirá un enfoque de investigación clásico [Riveros y Rosas, 1985; Creswell, 2002] con énfasis en la producción de tecnologías [Sábato y Mackenzie, 1982]; identificando métodos, materiales y abordaje metodológico necesarios para desarrollar el proyecto:

Métodos:

Revisiones Sistemáticas:

Las revisiones sistemáticas [Argimón, 2004] de artículos científicos siguen un método explícito para resumir la información sobre determinado tema o problema. Se diferencia de las revisiones narrativas en que provienen de una pregunta estructurada y de un protocolo previamente realizado.

Prototipado Evolutivo Experimental (Método de la Ingeniería):

El prototipado evolutivo experimental [Basili, 1993] consiste en desarrollar una solución inicial para un determinado problema, generando su refinamiento de manera evolutiva por prueba de aplicación de dicha solución a casos de estudio (problemáticas) de complejidad creciente. El proceso de refinamiento concluye al estabilizarse el prototipo en evolución.

Materiales:

Para el desarrollo de los formalismos y procesos propuestos se utilizarán:

- Formalismos de modelado conceptual usuales en la Ingeniería de Software [Rumbaugh et al., 1999; Jacobson et al., 2013] y en la Ingeniería del Conocimiento [García-Martínez y Britos, 2004].

- Modelos de Proceso usuales en Ingeniería de Software [IEEE, 1997; ANSI/IEEE, 2007; Oktaba et al., 2007].
- Normativas de Mercados de Capitalesvigentes sobre las responsabilidades del Estado [Ley 26831 de Mercados de Capitales, 2012].

RESULTADOS ESPERADOS

Se espera tener las versiones de los siguientes productos: (i) Proceso de Construcción de Algoritmos Basados en Inteligencia Artificial, (ii) Modelo de reconocimiento de los patrones de errores comúnmente realizadas en el desarrollo de modelos de trading algorítmico, (iii) Proceso de Auditoria para la evaluación de los algoritmos utilizados en trading algorítmico.

REFERENCIAS

- Eric Allen. Bug Patterns In Java. Apress; Softcover reprint of the original 1st ed. Edition. ISBN-10: 1590590619. 2002.
- Joel Hasbrouck, Gideon Saar. Low-latency trading. Journal of Financial Markets. Volume 16, Issue 4, November 2013, Pages 646-679. <https://doi.org/10.1016/j.finmar.2013.05.003>
- Chordia, Tarun and Goyal, Amit and Lehmann, Bruce N. and Saar, Gideon, High-Frequency Trading (June 2013). Johnson School Research Paper Series No. #20-2013. Available at SSRN: <https://ssrn.com/abstract=2278347> or <http://dx.doi.org/10.2139/ssrn.2278347>
- Wijnand Nuij, Viorel Milea, Frederik Hogenboom, Flavius Frasinca, Uzay Kaymak. An Automated Framework for Incorporating News into Stock Trading Strategies. IEEE Transactions on Knowledge and Data Engineering (Volume: 26, Issue: 4, April 2014) DOI: 10.1109/TKDE.2013.133
- Cain Evans, Konstantinos Pappas, Fatos Xhafa. Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign exchange speculation. Mathematical and Computer Modelling Volume 58, Issues 5–6, September

- 2013, Pages 1249-1266. <https://doi.org/10.1016/j.mcm.2013.02.002>.
- Jan Kleinnijenhuis, Friederike Schultz, Dirk Oegema. Financial news and market panics in the age of high-frequency sentiment trading algorithms. *Journalism*. First Published February 10, 2013.
 - ALAIN P. CHABOUD, BENJAMIN CHIQUOINE, ERIK HJALMARSSON, CLARA VEGA Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. *Journal of Finance*. DOI: 10.1111/jofi.12186.
 - Ekkehart Boehmer, Kingsley Y. L. Fong, J. (Julie) Wu. International Evidence on Algorithmic Trading. AFA 2013 San Diego Meetings Paper. Date Written: September 17, 2013.
 - Martin Scholtus, Dickvan Dijk, Bart Frijns. Speed, algorithmic trading, and market quality around macroeconomic news announcements. *Journal of Banking & Finance* Volume 38, January 2014, Pages 89-105. <https://doi.org/10.1016/j.jbankfin.2013.09.016>
 - Jonathan Brogaard Terrence Hendershott Ryan Riordan. High-Frequency Trading and Price Discovery. *The Review of Financial Studies*, Volume 27, Issue 8, 1 August 2014, Pages 2267–2306, <https://doi.org/10.1093/rfs/hhu032>.
 - Jakob Arnoldi. Computer Algorithms, Market Manipulation and the Institutionalization of High Frequency Trading. *Journal of Theory, Culture & Society*. First Published February 26, 2015
 - Yong Hua, Kang Liu, Xiangzhou Zhanga, Lijun Sub, E.W.T. Ngaic, Mei Liu. Application of evolutionary computation for rule discovery in stock algorithmic trading: A literature review. *Applied Soft Computing* Volume 36, November 2015, Pages 534-551. <https://doi.org/10.1016/j.asoc.2015.07.008>.
 - Cooper, Ricky; Ong, Michael; Van Vliet, Ben. Multi-scale capability: A better approach to performance measurement for algorithmic trading. *Journal: Algorithmic Finance*, vol. 4, no. 1-2, pp. 53-68, 2015. DOI: 10.3233/AF-150043.
 - José Manuel Berutich, Francisco López, Francisco Luna, David Quintana. Robust technical trading strategies using GP for algorithmic portfolio selection. *Expert Systems with Applications* Volume 46, 15 March 2016, Pages 307-315. <https://doi.org/10.1016/j.eswa.2015.10.040>.
 - Jonathan Brogaard, Terrence Hendershott, Ryan Riordan. Price Discovery Without Trading: Evidence from Limit Orders. *Finance Down Under 2017 Building on the Best from the Cellars of Finance*. Date Written: March 2016.
 - Tina Viljoen, P. Joakim Westerholm, Hui Zheng Algorithmic Trading, Liquidity, and Price Discovery: An Intraday Analysis of the SPI 200 Futures. *The Financial Review*. DOI: 10.1111/fire.12034.

GA²LA: Grupo de Aprendizaje Automático, Lenguajes y Autómatas

Juan C. Vázquez¹, Julio Castillo¹, Leticia Constable¹, Marina Cardenas¹

Depto. de Ingeniería en Sistemas de Información / Facultad Regional Córdoba / Universidad Tecnológica Nacional

¹{jcjvazquez, jotacastillo, leticiaconstable, angelaesmeralda}@gmail.com

RESUMEN

Durante más de catorce años de desarrollo de proyectos de investigación en las áreas de autómatas, lenguajes y redes neuronales artificiales, entre otras muchas temáticas afines, se ha conformado una masa crítica de recursos humanos capacitados, dotados de los conocimientos, experiencia y trayectoria para generar un consistente equipo de trabajo. La necesidad de conformar un ámbito más estructurado, concentrado y formal donde se establezca la discusión científica, académica y práctica de los distintos temas involucrados, de las líneas de acción a emprender a futuro, de las formas de financiamiento a utilizar y del modo trabajo que promueva la sinergia de equipo, guían la propuesta de conformación de un Grupo UTN de I+D en temáticas afines a los proyectos ya realizados y en marcha de estos investigadores. El “*Grupo de Estudio, Investigación, Desarrollo y Transferencia en Aprendizaje Automático, Lenguajes y Autómatas*” (GA²LA) es presentado aquí con sus antecedentes, expectativas y líneas de investigación.

Palabras clave: redes neurales artificiales, autómatas, lenguaje formal, lenguaje natural.

CONTEXTO

En 2002, se crea el Laboratorio de Investigación de Software (LIS) en la Facultad Regional Córdoba de la Universidad Tecnológica Nacional, con aportes de Microsoft Research de E.E.U.U. y de la Universidad. Las actividades de investigación y desarrollo inician formalmente en el ámbito académico, aproximadamente por el año 2004. Desde entonces los sucesivos integrantes del LIS han definido y llevado adelante proyectos de investigación acreditados por la Secretaría de Cien-

cia y Técnica de la UTN y otros organismos (Agencia Córdoba Ciencia, Ministerio de Ciencia y Tecnología de la Prov. de Córdoba, CONICET, ANPCYT), relacionados al estudio, diseño, desarrollo, implementación y transferencia, en las temáticas de ingeniería de software, modelos evolutivos y conexionistas, máquinas abstractas/autómatas, procesamiento de lenguaje natural y de lenguajes formales. Estos proyectos que se encuadran en líneas afines de investigación son los que dan origen a la idea de conformar el grupo que presentamos en este artículo.

En UTN, “Grupo” y “Centro” son entidades formales de I+D con características especiales definidas en sus estatutos.

Algunos proyectos de los integrantes y exintegrantes del LIS han formado parte durante años del Grupo de Inteligencia Artificial (GIA) desde su constitución en 2007 y del CIDS (Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información, ex grupo GIDTSI) desde su creación en 2014 (entidades de las cuales el director propuesto para el nuevo grupo GA²LA es miembro fundador). Así éstos, se constituyen en el origen natural del nuevo grupo en formación.

Es oportuno señalar, que quien se propone como director del grupo GA²LA fue el primer jefe a cargo del LIS y el codirector propuesto es el actual jefe a cargo del LIS, por lo que seguimos en contacto con el mismo.

1. INTRODUCCIÓN

Como producto de las actividades de I+D indicadas, en la temática específica del grupo GA²LA se desarrollan y se han desarrollado los siguientes proyectos:

a. Relaciones entre los momentos de aprendizaje y de reconocimiento de las redes neuronales artificiales y la evolución espacio-temporal de autómatas celulares (2004-2005): este proyecto estudió la historia de los pesos sinápticos de una red neuronal artificial tipo perceptrón multicapa durante su entrenamiento, intentando determinar si la misma podía ser construida por algún autómata celular del tipo unidimensional. También se estudió la historia de los estados de las neuronas de redes Hopfield durante la fase de reconocimiento de patrones, con el mismo fin.

b. Construcción de herramientas didácticas para la enseñanza y ejercitación práctica en laboratorio de la asignatura Sintaxis y Semántica del Lenguaje en la carrera de Ingeniería en Sistemas de Información (2006-2008): este proyecto sistematizó el conocimiento previsto para la asignatura SSL, generando una herramienta de autoevaluación y simuladores de máquinas abstractas y procedimientos sobre gramáticas formales.

c. Evaluación de arquitecturas y técnicas de entrenamiento de redes neuronales multicapa de perceptrones (2006-2008): su intención fue determinar cómo lograr el afinado del conjunto de parámetros que definen una red neuronal artificial, para obtener el mejor desempeño posible de la misma en sus tareas de reconocimiento.

d. Redes neuronales artificiales y autómatas celulares. Estudio y aplicaciones (2008-2011): seguido de (a), además de seguir con su objetivo inicial, aquí se desarrollaron aplicaciones de software bajo convenios, para empresas de desarrollo de software (VATES S.A.), para reconocimiento de imágenes de preparados de laboratorio de la Facultad de Veterinaria de la UCC y para el Instituto de Investigaciones Geo-Históricas (IIGHI-CONICET) sobre la metodología para el cálculo del riesgo para la salud de las viviendas urbanas. Este último fue además transferido a centros de investigación en Argentina, Paraguay, Brasil, Cuba y Colombia.

e. Transferencia del sistema de software RVS 2.1 para el cálculo de riesgo de la vivienda saludable en la Provincia de Córdoba (2008-

2009): con subsidio del Gobierno de Córdoba, se transfirieron productos de software y metodología diagnóstica a la Municipalidad de Villa del Totoral, Córdoba, Argentina.

f. El riesgo de la vivienda urbana para la salud desde la perspectiva de la salud comunitaria y la vulnerabilidad sociodemográfica en las ciudades Resistencia (Chaco) y Córdoba (2009-2010): este es un proyecto PICT del IIGHI-CONICET del cual se participó como parte del grupo responsable, para reforzar las transferencias realizadas.

g. Modelado para la predicción de incendios forestales en la Provincia de Córdoba (2010-2012): utilizando datos históricos e índices específicos de incendios, se desarrolló un modelo y herramienta para predecir dónde y de qué envergadura podrían producirse incendios en la zona serrana de Córdoba.

h. Aseguramiento de la trazabilidad en proyectos de desarrollo de sistemas de software (2010-2014): se estudió la aplicabilidad de los procesos de trazabilidad a los distintos tipos de sistemas desarrollados en distintos tipos de empresas. Se condujeron encuestas en empresas del medio para obtener datos históricos que, además de métodos estadísticos que fueron utilizados, alimentaron una red neural para efectuar predicciones sobre trazabilidad.

i. Redes neuronales artificiales y autómatas celulares. Productos y aplicaciones (2012-2014): seguido desde (d), este proyecto actualizó las herramientas de software construidas en anteriores proyectos y realizó difusión y divulgación, en conjunto con investigadores de IIGHI-CONICET. Se realizó una transferencia a la Municipalidad de Benevides en el estado de Belem-Pará, Brasil y se estudian nuevas transferencias.

j. Construcción de un modelo de pronósticos para predicción de incendios forestales en la Provincia de Córdoba, (2013-2016): seguido de (g), se profundizó el estudio del modelo, se generaron herramientas para obtención de información histórica por Internet y se entrenaron nuevas redes neuronales para efectuar predicciones.

k. Desarrollo de un sistema de análisis de tex-

to no estructurado (2012-2013) / Fase II (2014-2015) / Fase III (2016-2017/18): estos proyectos (en relación con la tesis doctoral de uno de los integrantes de GA²LA), mediante técnicas de inteligencia artificial para el procesamiento de textos escritos en español e inglés, intenta determinar implicaciones lógicas entre distintas partes del mismo o entre distintos textos.

l. Evaluación del impacto de variantes no convencionales en el desempeño de autómatas finitos con memoria de pila (2015-2017/18): se definieron máquinas con memoria de pila con accesos no usuales, para determinar su efectividad en la resolución de distintos tipos de problemas. Se construyeron simuladores de estas máquinas que ayudaron en la experimentación.

m. Detección de errores sintácticos bajo el algoritmo de Earley (2014-2017/18): aquí se intenta determinar cómo se pueden informar adecuadamente (localización y tipo) sobre errores sintácticos de sentencias analizadas con el algoritmo de Earley. Se desarrollaron analizadores léxicos y sintácticos para experimentación y se estudia la relación entre el funcionamiento del algoritmo Earley y las redes de Petri, entre muchos otros aspectos.

n. Redes neurales artificiales con aprendizaje profundo (2018-2020): este es un nuevo proyecto iniciado muy recientemente, que se interesa primeramente por la parte teórica del aprendizaje profundo para establecer cómo se visualizan las características de distinto nivel que esta técnica extrae de los datos de entrenamiento. Por otro lado, trabajará con las herramientas disponibles y posiblemente genere nuevas, para actualizar las transferencias de redes neurales entrenadas efectuados en anteriores proyectos.

En todos estos proyectos, además de los logros específicos alcanzados según sus objetivos de investigación y desarrollo, se formaron numerosos recursos humanos de grado (ya sea mediante prácticas supervisadas de alumnos próximos a recibirse, como a becarios alumnos) y de posgrado (becarios graduados, especialidades, maestrías y doctorados).

Se efectuaron además numerosas transferencias tanto a centros de investigación, como a empresas o entes oficiales de gobierno.

Por otro lado, los integrantes de estos proyectos se desempeñan en numerosas cátedras de distintas carreras de la Facultad Córdoba de la UTN y de otras universidades (Nacional de Córdoba, Instituto Universitario Aeronáutico, Blas Pascal, Católica de Córdoba, Nacional de Villa María) y en carreras de posgrado de las mismas. Por ello, siempre parte de los proyectos contienen algún objetivo académico, por lo cual a las transferencias anteriores se deben agregar las transferencia efectuadas a las mismas cátedras de las cuales son parte los investigadores, mediante artículos, conferencias y seminarios.

Finalmente, todos los proyectos han realizado difusión de los conocimientos logrados en los más diversos congresos nacionales e internacionales de informática y otros relacionados, participación en la redacción de libros y capítulos de libros, y también se ha realizado divulgación de las investigaciones en medios periodísticos, charlas en escuelas de nivel medio y artículos en páginas web.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

De los proyectos indicados en la Introducción, se desprenden claramente al menos tres líneas de investigación y desarrollo interrelacionadas que el grupo GA²LA tendrá como ejes para su funcionamiento, a saber:

- Autómatas y Lenguajes Formales
- Procesamiento de Lenguaje Natural
- Aprendizaje Automático

Dentro del grupo propuesto, sus integrantes participan de entre muchas, de las cátedras de Sintaxis y Semántica de los Lenguajes, Investigación Operativa, Inteligencia Artificial, Algoritmos y Estructuras de Datos, Paradigmas de Programación, Simulación, Matemática Discreta y Análisis Matemático, por lo que las líneas se mezclan “adecuadamente” con sus inquietudes y trabajo académico habitual.

Además, han efectuado o tienen en marcha, posgrados relacionados con las mismas.

3. RESULTADOS OBTENIDOS / ESPERADOS

Los integrantes del grupo GA²LA en conjunto, como resultado de los proyectos realizados entre otras cosas, a la fecha han:

- Realizado más de 70 presentaciones y publicaciones en congresos nacionales e internacionales.
- Obtenido derechos de autor por productos de software en cuatro oportunidades.
- Escrito 11 libros o capítulos de libros.
- Participado en la organización de al menos siete congresos y workshops nacionales.
- Participado en la evaluación de: artículos en congresos, revistas y journals nacionales e internacionales, proyectos en el Ministerio de Ciencia y Tecnología de Córdoba, categorización de investigadores en UTN, competencias de programación y de desarrollo de software, etc.
- Participado de numerosas transferencias bajo convenio a centros de investigación, empresas y entes de gobierno local.
- Dirigido tesis de maestría, especialidad, numerosas prácticas supervisadas y becarios tanto de grado como graduados.

Se espera que trabajando organizados en un grupo formal, estos logros puedan no solo mantenerse en el tiempo, sino incrementarse.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo que se presenta está conformado por seis investigadores formados, seis investigadores noveles y alumnos de grado y graduados (como becarios) de la Universidad Tecnológica Nacional, Facultad Regional Córdoba.

Actualmente, en los proyectos en proceso, se tienen en formación:

- Dos doctorandos

- Dos maestrandos
- Un becario graduado
- Siete becarios de grado

Se piensa sumar en el corto plazo un nuevo doctorando y becarios de grado para los proyectos vigentes durante 2018.

5. BIBLIOGRAFÍA

1. Castillo, J., Cardenas, M., Medel, R., Casco, O., Navarro, M. & Gutierrez, S. "SISTEMA PARA PREDICCIÓN DE INCENDIOS APLICADO A LA PROVINCIA DE CÓRDOBA". (2016). XVII Workshop de Investigadores en Ciencias de la Computación, WICC 2016, Entre Ríos, Argentina.
2. Castillo, Julio J.; Cardenas, Marina E.; Curti, Adrián; Casco, Osvaldo. "Software para asistencia en la creación de corpus para sistemas de análisis de texto no estructurado". WICC 2015. XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015).
3. Marina Cardenas, Juan C. Vázquez, Julio Castillo, Sebastián Villena Ruiz. "SISTEMA DE PREDICCIÓN DE INCENDIOS FORESTALES BASADO EN EL ÍNDICE FWI PARA LA PROVINCIA DE CÓRDOBA". XV Workshop de Investigadores en Ciencias de la Computación. WICC 2013. Paraná, Entre Ríos, Argentina. Mayo de 2013.
4. Juan C. Vázquez, Julio Javier Castillo, Marina E. Cardenas, Romina Gordillo, Sebastián Villena Ruiz. "Predicción de Incendios Forestales en la Provincia de Córdoba". Lugar: Posadas, Misiones, Argentina. Mayo de 2012. WICC 2012.
5. Juan C. Vázquez, Julio Javier Castillo, Marina E. Cardenas, María del Carmen Rojas. "Modelo Computacional empleando Redes Neuronales Artificiales para la estimación del Riesgo para la Salud de la Vivienda Urbana". WICC 2011 – XIII Workshop de Investigadores en Ciencias de la Computación. Mayo de 2011. Fecha: 5 Y 6 DE

- MAYO DE 2011 Rosario, SANTA FE, ARGENTINA.
6. Juan C. Vázquez, Julio Castillo, María del C. Rojas, Marina Cardenas. “Modelización y Estimación del Riesgo para la Salud de la Vivienda Urbana empleando Redes Neuronales Artificiales”, WICC 2010, Mayo de 2010. ISBN: 978-950-34-0652-6. 5 Y 6 DE MAYO DE 2010 CALAFATE, SANTA CRUZ, ARGENTINA.
 7. Vázquez Juan, Castillo Julio J., Rojas M. del Carmen, Marciszack Marcelo. Redes Neuronales Artificiales aplicadas a Ciencias Sociales. WICC 2008 - X Workshop de Investigadores de Ciencias de la Computación. Entidad Responsable: Red UNCI. Ciudad: General Pico, La Pampa, Argentina. Idioma: Español. ISBN: 978-950-863-101-5.
 8. Una estrategia didáctica para el dictado de Informática Teórica; X Workshop de Investigadores en Ciencias de la Computación (WICC 2008); 2008; General Pico, La Pampa, Argentina., Autores: Marciszack M., Cárdenas M., Vázquez J., Castillo J., ISBN: 978-950-863-101-5 (CD)
 9. Construcción de herramientas didácticas para la enseñanza y la ejercitación práctica en laboratorio de “Informática Teórica” en las carreras con Informática; VIII Workshop de Investigadores en Ciencias de la Computación (WICC 2006); Junio de 2006; Morón, Buenos Aires, Argentina. Autores: Marciszack M., Vázquez J., Serrano D., Pérez R., ISBN: 950-9474-35-5, ISBN13: 978-950-9474-35-2 (CD), ISBN: 950-9474-34-7, ISBN13: 978-950-9474-34-5 (Impreso)
 10. Detección de errores sintácticos bajo el algoritmo Earley; XVII Workshop de Investigadores en Ciencias de la Computación (WICC 2015); Abril de 2015; Salta, Salta, Argentina. Autores: Vázquez J., Constable L., Brenda M., Jornet W., Arcidiacono M., Parisi G., ISBN: 978-987-633-134-0 – Red UNCI/UNSA
 11. Tratamiento de Imágenes de Fibras Animales; Constable L., Arcidiacono M., Vázquez J., Picco J.; XV Workshop de Investigadores en Ciencias de la Computación (WICC 2013); Red UNCI, Universidad Autónoma de Entre Ríos, Facultad de Ciencia y Tecnología; Paraná, Entre Ríos, Argentina; 2013. ISBN: 978-987-28179-6-1 (CD)
 12. Verificación estática de confidencialidad en un sistema de múltiples niveles de seguridad basado en Java bytecodes; Medel R., Spessot C., Vázquez J., Giagante I.; XI Workshop de Investigadores en Ciencias de la Computación (WICC 2009), Universidad Nacional de San Juan; San Juan, San Juan, Argentina; 2009. ISBN: 978-95060-557-0-7 (CD)
 13. Redes Neuronales Artificiales aplicadas a Ciencias Sociales; Vázquez J., Castillo J.; X Workshop de Investigadores en Ciencias de la Computación (WICC 2008), Universidad Nacional de La Pampa; General Pico, La Pampa, Argentina. Mayo de 2008. ISBN: 978-950-863-101-5 (CD)
 14. Redes Neuronales Artificiales vs. Automatas Celulares; Martínez F., Vázquez J., Marciszack M.; VII Workshop de Investigadores en Ciencias de la Computación (WICC 2005); Río Cuarto, Córdoba, Argentina. Mayo de 2005. ISBN: 950-665-337-2 (CD)

Sistemas de análisis textual en formato no estructurado

Julio Castillo¹, Marina Cardenas¹, Martin Navarro¹,
Nicolas Hernandez¹, Melisa Velazco¹

¹ Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/ Facultad Regional Córdoba/ Universidad Tecnológica Nacional
{jotacastillo, ing.marinacardenas}@gmail.com

Resumen

En este artículo se describen las actividades desarrolladas y los subsistemas que conforman el proyecto de investigación denominado *Desarrollo de Sistemas de Análisis de Texto*.

Este proyecto aborda la problemática del desarrollo de herramientas que permitan recolectar, tabular, y etiquetar, textos en diferentes formatos y de diferentes fuentes de información con el propósito de someterlos a un posterior análisis utilizando aprendizaje automático y técnicas de minería de datos. Además del desarrollo de estas herramientas, el proyecto contempla el desarrollo de sistemas de análisis de texto que puedan abordar problemas como el reconocimiento de paráfrasis, es decir identificar oraciones (o párrafos) que tengan el mismo significado, o bien identificar oraciones-párrafos que estén semánticamente relacionados entre sí mediante una relación de implicación.

Las líneas de investigación en la que se encuadra el proyecto es dentro de las áreas de lingüística computacional y de aprendizaje automático. En particular, el proyecto se enfoca en modelos que utilizan redes neuronales artificiales (RNA) para analizar y procesar textos no estructurados.

Palabras clave: análisis de texto, extracción de información, corpus, machine learning.

Contexto

El presente proyecto denominado Análisis de Texto (ADT) es un proyecto homologado por la SCyT de la UTN, y se enmarca dentro del área de computación

lingüística. El mismo se desarrolla en el Laboratorio de Investigación de Software LIS¹ del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN-FRC).

Actualmente, el proyecto se encuentra dentro del grupo de investigación denominado Grupo de Inteligencia Artificial (o GIA) de la UTN-FRC.

Este grupo GIA nuclea proyectos de una línea de investigación relacionada al área de inteligencia artificial, redes neuronales artificiales, análisis y procesamiento de imágenes.

El grupo se conforma por doctores, ingenieros, licenciados, becarios y pasantes.

El proyecto de análisis de texto, junto con otros proyectos, la mayoría de los ellos surgidos en el laboratorio de investigación de software de la UTN-FRC, han dado origen a varias líneas de investigación consolidadas (teoría de autómatas y gramáticas formales, modelos de predicciones de fenómenos climatológicos, entre otros). Esta sinergia entre múltiples proyectos y líneas de investigación han llevado a la necesidad de creación de un nuevo grupo de investigación UTN el cual se encuentra en una etapa de formación, tal como se detalla en la sección de líneas de investigación, desarrollo e innovación.

1. Introducción

El proyecto denominado Desarrollo de Sistemas Análisis de Texto (ADT), aborda

¹ www.investigacion.frc.utn.edu.ar/mslabs/

dos grandes problemas claramente diferenciados. El primero, relacionado a la necesidad de obtener información para poder construir corpus lingüísticos. El segundo, relacionado con el desarrollo de sistemas que hagan uso de dichos corpus para encarar problemas complejos del lenguaje natural.

En este sentido, el proyecto aborda el problema del análisis e interpretación de textos no estructurados, extracción de información y minería de datos [1][2][3][4][5] basados en técnicas de aprendizaje automático por computadora, especialmente aquellas basadas en redes neuronales artificiales [6][7][8], máquinas kernel [9], deep learning [10][11], y árboles de decisión entre otras.

En el marco de este proyecto se han desarrollado, y se continúan desarrollando varios sistemas de análisis y procesamiento de texto, entre los que se mencionan:

- Software de Asistente de Creación de Corpus (ACC): es un software que permite construir material de entrenamiento para aplicaciones de minería de datos sobre texto no estructurado.
- Sistema de Mapeo de Datos (SMD): Software que permite manipular orígenes de datos estructurados y centralizarlos para un posterior análisis con técnicas de recuperación de información o de minería de datos.
- Sistema de detección de similitudes en archivos de código fuente (SDS). Es un sistema que se está comenzando a desarrollar y que tiene como objetivo analizar archivos de código fuente escritos en diferentes lenguajes de programación e informar el grado de similitud entre los mismos.

2. Líneas de Investigación, Desarrollo e Innovación

La línea de investigación principal y específica de este proyecto es el abordaje de problemáticas de lingüística computacional utilizando aprendizaje automático.

De esto subyace la línea demarcada por la Lingüística de Corpus [12], entendida como el estudio empírico de la lengua a partir de los datos que proporcionan ejemplos reales de producciones lingüísticas (orales o escritas) almacenadas en una computadora.

Esta línea abarca un campo científico interdisciplinar cuyo principal objetivo es el de desarrollar sistemas con la capacidad de reconocer y comprender el lenguaje natural humano a través de modelos computacionales.

Muy emparentada a esta línea de investigación se encuentra otra línea de investigación relacionada a la construcción de modelos computacionales de predicción de incendios forestales y de fenómenos climatológicos. En esta línea también participan integrantes de este proyecto.

La confluencia del trabajo de varias líneas de investigación (teoría de autómatas y gramáticas formales, modelos de predicciones de fenómenos climatológicos, y el modelado de problemas del área de ciencias sociales), consolidadas en el tiempo han llevado a la creación formal de un nuevo grupo UTN de investigación. Actualmente, la aprobación de dicho grupo se encuentra en trámite dentro de la UTN.

3. Resultados

En el proyecto se han desarrollado varios sistemas de análisis y procesamiento de texto, entre los más importantes mencionaremos a un sistema Software de Asistente de Creación de Corpus (ACC), un Sistema de Mapeo de Datos (SMD), y un Sistema de detección de similitudes en

archivos de código fuente (SDS). Estos sistemas se describen con más detalle a continuación.

El Software de Asistente de Creación de Corpus (ACC) se desarrolla con el objetivo de facilitar la construcción de material de entrenamiento que se necesita en los algoritmos de aprendizaje supervisado. La calidad y el tamaño del conjunto de entrenamiento impacta directamente en la efectividad de los algoritmos de clasificación, es por ello que se necesita un tamaño adecuado del material de entrenamiento y que el mismo sea consistente.

Adicionalmente, el ACC permite registrar diversos fenómenos lingüísticos a nivel léxico, sintáctico, morfológico y semántico. Se trata de un software que ya está desarrollado pero al cual se le continúan agregando nuevas funcionalidades. Entre los resultados logrados por este asistente podemos destacar:

- Lectura de corpus: Se realiza la lectura de corpus del NIST (National Institute of Standards and Technology) para su posterior generación, tabulación, ordenamiento y etiquetado, como así también la traducción del material al español utilizando el traductor automático de Google Translate.

- Carga de pares del corpus.

- Búsqueda y posicionamiento de un par dentro del corpus.

- Selección de subcadenas de fragmentos de texto para someterlos a una posterior clasificación: esto permite seleccionar partes de un texto y visualizarlas gráficamente a través de una tabla para su posterior modificación.

- Clasificación de los fenómenos en categorías y subcategorías.

- Creación de un nuevo corpus etiquetado almacenado en formato .xml .

Por otra parte, el programa permite acelerar el tiempo necesario para la confección del material de entrenamiento, como así también brinda trazabilidad respecto de los expertos humanos que contribuyeron a cada parte del corpus. Esto permite establecer métricas y calcular la confianza del material de entrenamiento construido.

Entre las aplicaciones que potencialmente podrían utilizar este material de entrenamiento podemos citar a traducción automática asistida por computador, creación de corpus de paráfrasis, creación de corpus para implicación de textos, resumen automático, entre otras posibles aplicaciones.

A la fecha se han creado tres corpus monolingües. Un corpus consta de 50 pares en inglés, los otros dos corpus están en español, cada uno presenta 100 pares de elementos etiquetados con información lingüística, de acuerdo a en los que se describe en [13] y con las clasificaciones enumeradas en [14].

Como herramienta, el ACC ha contribuido a los objetivos del proyecto proveyendo de material de entrenamiento tanto en el idioma español, como en el inglés. Esto ha facilitado y mejorado el funcionamiento de los Sistemas de RTE (Implicación Textual).

Actualmente, el ACC se sigue utilizando en la generación de corpus, y se está estudiando la posibilidad de dejarla disponible para el acceso libre de otros investigadores que deseen hacer uso de la misma en sus trabajos.

El software de Sistema de Mapeo de Datos (SMD) se plantea con el objetivo de realizar una manipulación, procesamiento (desde diferentes fuentes y orígenes de datos), y almacenamiento de la información en un repositorio común centralizado (una base de datos en SQL Server). Se pretende entonces, explotar el repositorio con diversas técnicas del área de minería de

datos y técnicas de recuperación de la información.

Hay que notar, que este sistema necesita mantenerse actualizado para que la información del repositorio sea correcta y fiable. El lapso de tiempo necesario entre cada actualización dependerá de la aplicación que se esté desarrollando.

Finalmente, mencionamos el sistema de detección de similitudes en códigos fuente (SDS). Este software está en sus primeras etapas y presenta como objetivo el permitir cuantificar el grado de similitud entre dos archivos de texto plano, en particular dos archivos de código fuente escritos en el mismo lenguaje de programación.

Hasta el momento, se ha desarrollado un prototipo que incluye una simple interfaz gráfica en la que es posible seleccionar dos archivos de códigos fuentes para aplicarles medidas principalmente de similitud léxica, y se está integrando una medida de similitud sintáctica.

El SDS también permite la comparación del tipo 1-N, en la cual un código fuente es comparado con N-elementos de un conjunto. Esta operación demanda un alto costo computacional, por lo que se están estudiando y desarrollando técnicas basadas en paralelismo para poder encontrar las similitudes de todos los elementos de un conjunto e informarlas de manera ordenada en función del grado de similitud.

En este caso, una detección de similitud que contemple a todos los elementos de un conjunto de N-elementos requerirá $(N*N)/2$ comparaciones entre los archivos fuente. Es por ello, la necesidad de contar con algoritmos eficientes de detección de similitud.

4. Formación de Recursos Humanos

El equipo de investigación y desarrollo de software, está formado por docentes investigadores de la Universidad

Tecnológica Nacional, Facultad Regional Córdoba, que a continuación se detallan:

- Un doctor en ciencias de la computación, quién guía a becarios de grado y de posgrado, como así también realiza la dirección de prácticas profesionales supervisadas y pasantías.
- Una magister en ingeniería en sistemas de información que está en la etapa de escritura del plan de tesis para iniciar su doctorado en ingeniería en las temáticas de la línea de investigación mencionada. La unidad académica de radicación del doctorando sería la Universidad Tecnológica Nacional, Facultad Regional Córdoba. También realiza la dirección de becarios de posgrado y de becarios de grado en el contexto del presente proyecto.
- Anualmente participan en el proyecto, entre dos y cuatro alumnos realizando su práctica supervisada, la cual es necesaria como parte de los requisitos para la obtención del grado de Ingeniero.
- El proyecto además posee investigadores en formación y en proceso de categorización.
- Año tras año se capacita y forma a alumnos becarios que participan en el proyecto y que realizan actividades de investigación, complementando de esta manera su formación curricular desde el punto de vista científico.
- Finalmente, se han realizado charlas de difusión y jornadas de capacitación a alumnos y a docentes de ingeniería en sistemas de información en las líneas temáticas enumeradas anteriormente.

5. BIBLIOGRAFÍA

[1] Judith Klavans y Philip Resnik. The Balancing Act. Combining Symbolic and Statistical Approaches to Language. MIT Press, 1996.

[2] C. Manning y H. Schutze. Foundations of Statistical Natural Language

Processing. The MIT Press, Cambridge, MA, 1999.

[3] Castillo J. Sagan in TAC2009: Using Support Vector Machines in Recognizing Textual Entailment and TE Search Pilot task. TAC, 2009.

[4] Castillo J., Cardenas M. Using Sentence Semantic Similarity Based on WordNet in Recognizing Textual Entailment. Iberamia 2010, LNCS, vol. 6433, pp. 366-375, 2010.

[5] Castillo J. Using Machine Translation Systems to Expand a Corpus in Textual Entailment. Proceedings of the Ictel 2010, LNCS, vol. 6233, pp.97-102, 2010.

[6] Feldman R. y Hirsh H.. Exploiting Background Information in Knowledge Discovery from Text. Journal of Intelligent Information Systems, 1996.

[7] Lewis, D.. Evaluating and optimizing autonomous text classification systems. In Proceedings of SIGIR-95, 18th ACM International Conference on Research and Development in Information Retrieval. Seattle, US, págs. 246-254, 1995.

[8] M. Craven y J. Shavlik. Using Neural Networks for Data Mining. Future Generation Computer Systems, 13, págs. 211-229, 1997.

[9] Castillo J. An approach to Recognizing Textual Entailment and TE Search Task using SVM. Procesamiento del Lenguaje Natural 44, 139-145, 2010. 4, 2010.

[10] I. Goodfellow, Y. Bengio y A. Courville. Deep Learning. MIT Press. 2016.

[11] N. Buduma. *Fundamentals of Deep Learning: Designing Next-Generation Artificial Intelligence Algorithms*. O'Reilly book. 2015.

[12] Stefan Th. Y Anatol Stefanowitsch. *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*, Berlin: Mouton, pág. 117, 2006.

[13] Castillo Julio J., Cardenas Marina E., Curti Adrián, Casco Osvaldo, Navarro Martín, Hernández Nicolás A., Velazco Melisa. Desarrollo de sistemas de análisis de texto. XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017). 2017.

[14] Castillo Julio, Cardenas Marina, Curti Adrian, Velazco Melisa, Casco Osvaldo, Navarro Martin. Herramientas para Aplicaciones de Análisis de Textos. Congreso Nacional de Ingeniería Informática / Sistemas de Información. CONAISI 2017.

Procesamiento de textos estructurados

Marina Cardenas, Julio Castillo

Laboratorio de Investigación de Software LIS, Dpto. Ingeniería en Sistemas de Información
Facultad Regional Córdoba, Universidad Tecnológica Nacional

{ing.marinacardenas, jotacastillo}@gmail.com

RESUMEN

En este artículo se describe un proyecto de investigación relacionado al procesamiento de textos estructurados, en particular a archivos de códigos fuentes en algún lenguaje de programación.

Se propone abordar este problema obteniendo información a diferentes niveles de abstracción, adaptando técnicas que son específicas de implicación textual.

El proyecto se inserta en una línea de investigación de aprendizaje automático y de lingüística computacional.

Se describe el proyecto, los resultados obtenidos hasta el momento, como así también, los resultados que se esperan de la utilización del sistema. Finalmente, se presenta la línea de investigación en la que se enmarca este proyecto.

Palabras clave: análisis de texto, extracción de información, corpus.

CONTEXTO

El proyecto acreditado por la Universidad Tecnológica Nacional denominado *Modelado para el procesamiento de textos estructurados* (código UTN4518) se encuentra consolidado dentro de la línea de investigación relacionada con lingüística computacional y es llevado a cabo en el Laboratorio de Investigación de Software LIS1 del Departamento de Ingeniería en

¹ www.investigacion.frc.utn.edu.ar/mslabs/

Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba.

A su vez, este proyecto se encuentra dentro del grupo de investigación denominado Grupo de Inteligencia Artificial (o GIA) de la UTN-FRC.

El grupo GIA reúne a proyectos de investigación que se hayan todos en temáticas concernientes a la inteligencia artificial entre las que podemos destacar análisis de imágenes, algoritmos evolutivos, y su aplicabilidad en problemas de la ingeniería, de las ciencias naturales, y de las ciencias sociales.

Este grupo está compuesto de becarios, pasantes, docentes investigadores y doctores.

1. INTRODUCCIÓN

En el proyecto de procesamiento de textos estructurados se plantea el desarrollo de un modelo para detección de similitudes de código fuente para poder determinar la existencia de prácticas de reutilización aplicando técnicas vinculadas a la lingüística computacional, tales como minería de datos sobre texto y procesamiento del lenguaje natural, dado que según [1] “los lenguajes de programación se parecen a los lenguajes naturales en tanto que ambos, códigos fuente y textos escritos en lenguaje natural, se pueden representar como cadenas de símbolos (caracteres, palabras, etc.)”.

La identificación de similitudes de código puede servir para varios propósitos [2], entre los que se puede mencionar el

estudio de la evolución del código fuente de un proyecto, detección de prácticas de plagio, detección de prácticas de reutilización, extracción de un fragmento de código para “refactorización” del mismo y seguimiento de defectos para su corrección.

La detección de similitudes de código fuente con fines de reutilización es una tarea laboriosa que puede demandar altos costos de tiempo y dinero, dependiendo del impacto de la tarea de detección según el ámbito de aplicación.

Actualmente existen diferentes aproximaciones para abordar la problemática planteada en el presente proyecto, alguna de ellas se mencionan a continuación:

- Aproximaciones basadas en atributos, donde las métricas se calculan a partir del código fuente y se utilizan para la comparación de los distintos archivos. Un ejemplo sencillo sería: utilizar el tamaño del código fuente (número de caracteres, palabras y líneas) como atributo comparable de tamaño [3], o también el número de variables, el número de funciones, el número de clases, entre otros.
- Aproximaciones basadas en Tokens, en las cuales se convierte el código fuente en una secuencia de “tokens” o marcas, para una posterior evaluación y selección de estas secuencias de tokens según ciertas métricas [4] [5].
- Aproximaciones basadas en la estructura, donde el código fuente es convertido a una representación intermedia interna (IR), la cual es utilizada posteriormente para la comparación [6] [7].

Dentro de este contexto, se propone la construcción de un sistema, basado en un modelo que utiliza técnicas de Aprendizaje Automático Supervisado, comúnmente utilizadas para minería de datos [8], que permita la detección de similitudes de código fuente en los lenguajes de programación Java y Python en base a un corpus que será elaborado especialmente para tal fin.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación en las que se enmarca el proyecto de modelado para el procesamiento de textos estructurados es el área de inteligencia artificial, más concretamente una sub-especialidad que se denomina computación lingüística. En particular, nos enfocamos en los enfoques basados en aprendizaje automático.

Los desarrollos de esta línea de investigación, lo constituyen, por un lado, las herramientas elaboradas para facilitar el análisis y procesamiento de archivos de textos, en este caso de código fuente, y por el otro, los sistemas de reconocimiento de similitudes entre dos archivos de código fuente.

La innovación del proyecto concierne a los nuevos métodos propuestos para el análisis y procesamiento de textos, como así también a los algoritmos creados para abordar las problemáticas anteriormente mencionadas. Los algoritmos diseñados aprovechan las diferentes características que se pueden aprender de los textos y que son recolectados y creados a partir de las herramientas de procesamiento de textos.

Son múltiples las posibles sub-disciplinas que podrían valerse de los resultados de este proyecto, entre las que podemos destacar a las tareas de recuperación de información, evaluación de las traducciones automáticas [9], evaluación de la calidad de las traducciones, reconocimiento de paráfrasis [10] e implicación de textos [11][12][13][14][15]. Adicionalmente, la creación de corpus sobre texto estructurado es una actividad de relevancia y que puede impactar en otras tareas relacionadas al procesamiento del lenguaje.

La confluencia del trabajo de varias líneas de investigación que se desarrollan en el Laboratorio de investigación de software LIS, lugar donde se lleva a cabo este proyecto, ha llevado a la necesidad de plantear el surgimiento de un nuevo grupo

UTN que nuclea a las siguientes líneas de investigación: teoría de autómatas y gramáticas formales, modelos de predicciones de fenómenos climatológicos, y el modelado de problemas del área de ciencias sociales.

En este contexto, el presente proyecto se haya incluido en la presentación de este nuevo grupo de investigación en UTN. Actualmente, la aprobación de dicho grupo se encuentra en trámite dentro de la UTN.

3. RESULTADOS OBTENIDOS/ESPERADOS

Este proyecto se encuentra en sus primeras etapas y, hasta el momento, se han conseguido los siguientes resultados.

Se ha desarrollado un prototipo que permite seleccionar dos archivos de códigos fuentes de Java o Python, sobre los que se puede aplicar un conjunto de medidas de similitud léxica. A su vez, se está integrando una medida de similitud sintáctica.

Hasta el momento, el sistema puede identificar dos archivos con alto grado de superposición léxica, pero no es capaz de identificar archivos similares que utilizan estructuras sintácticas equivalentes pero que se escriben utilizando sintaxis diferente. Claramente, se debe a la imposibilidad de capturar la información presente a un nivel más profundo que el nivel léxico.

Uno de los inconvenientes que se presentan actualmente es el tiempo necesario para la ejecución de las métricas sobre un conjunto de archivos. Supongamos que se desee obtener los dos archivos más similares de un conjunto. Para ello es necesaria una comparación entre cada par de elementos del conjunto, la cual es una operación muy costosa desde el punto de vista computacional, y es por esto que los algoritmos propuestos como medidas de similitud no solo tienen que ser efectivos, sino también escalables y paralelizables.

Un sistema de detección de similitudes que contemple a todos los elementos de un

conjunto de N-elementos requerirá $(N*N)/2$ comparaciones entre los archivos de código fuente.

Por consiguiente, se están desarrollando algoritmos que sean capaces de modelar de manera precisa el fenómeno de similitud textual a diferentes niveles léxico, sintáctico y semántico, pero al mismo tiempo cumpliendo con el requerimiento no funcional de que debe tratarse de un algoritmo de baja complejidad computacional.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está formado por docentes investigadores del Laboratorio de Investigación de Software LIS² del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba, se detallan a continuación los responsables del proyecto:

- Una magister en ingeniería en sistemas de información que está evaluando la posibilidad de desarrollar su tema de tesis de doctorado en la esta temática con una variación del enfoque desde el punto de vista de los sistemas de Generación del Lenguaje Natural (NLG), y dirige a los integrantes miembros del equipo.
- Un doctor en ciencias de la computación que desarrolló su tesis de doctoral en la temática de implicación de textos y paráfrasis, y colaborara en la coordinación de becarios.
- Participan del proyecto alumnos que necesitan realizar su práctica supervisada que es uno de los requisitos para la obtención del grado de Ingeniero. Los alumnos que intervienen aprenden a realizar actividades de investigación, y cómo integrarse en un equipo existente.

² www.investigacion.frc.utn.edu.ar/mslabs/

- Por año participan uno o dos becarios alumnos a los que se les enseña como trabajar en un proyecto de investigación, y como llevar adelante actividades de investigación.
- Un becario de investigación de posgrado se ha incorporado recientemente, por lo cual el proyecto complementará su formación profesional.

5. BIBLIOGRAFÍA

- [1] Frantzeskou, G., MacDonell, S., Stamatatos, E., Gritzalis S. (2008). Examining the significance of high-level programming features in source code author classification. *The Journal of Systems and Software*, 81(3):447–460.
- [2] Smith, R. y Horwitz, S. (2009). Detecting and Measuring Similarity in Code Clones. *International Workshop on Software Clones (IWSC'09)*, pp. 28-34.
- [3] Wise M. (1992). Detection of similarities in student programs: YAP'ing may be preferable to plaguing. In *ACM SIGCSE Bulletin*, volume 24, pp. 268–271.
- [4] Wise M. (1993) Running Karp-Rabin matching and greedy string tiling. *Basser Dept. of Computer Science, University of Sydney, Sydney*.
- [5] Schleimer, S., Wilkerson, D., y Aiken, A. (2003). Winnowing: Local Algorithms for Document Fingerprinting. En: *Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data*, pp. 76-85.
- [6] Li, X., y Zhong, X. (2010). The source code plagiarism detection using AST. In *International Symposium IPTC*, pp. 406–408.
- [7] Baxter I., Yahin, A., Moura, L., Sant'Anna, M., y Bier, L. (1998). Clone detection using abstract syntax trees. En *Proceedings de IEEE ICSM 1998*, pp. 368–377.
- [8] Jadon, S. (2016). Code clones detection using machine learning technique: Support vector machine. *2016 International Conference on Computing, Communication and Automation (ICCCA)*. IEEE. Noida, India.
- [9] FeldmanR. y Hirsh H.. Exploiting Background Information in Knowledge Discovery from Text. *Journal of Intelligent Information Systems*, 1996.
- [10] Lewis, D.. Evaluating and optimizing autonomous text classification systems. In *Proceedings of SIGIR-95, 18th ACM International Conference on Research and Development in Information Retrieval*. Seattle, US, págs. 246-254, 1995.
- [11] M. Craven y J. Shavlik. Using Neural Networks for Data Mining. *Future Generation Computer Systems*, 13, págs. 211-229, 1997.
- [12] Castillo J. An approach to Recognizing Textual Entailment and TE Search Task using SVM. *Procesamiento del Lenguaje Natural* 44, 139-145, 2010. 4, 2010.
- [13] I. Goodfellow, Y. Bengio y A. Courville. *Deep Learning*. MIT Press. 2016.
- [14] Castillo J. Using Machine Translation Systems to Expand a Corpus in Textual Entailment. *Proceedings of the Iccetal 2010, LNCS*, vol. 6233, pp.97-102, 2010.
- [15] Castillo J., Cardenas M. Using Sentence Semantic Similarity Based on WordNet in Recognizing Textual Entailment. *Iberamia 2010, LNCS*, vol. 6433, pp. 366-375, 2010.

Arquitectura, Redes y Sistemas Operativos

Análisis de tráfico multicast de video H264/Theora en Redes Wi-Fi IEEE 802.11ac

Higinio Facchini, Santiago Pérez, Fabian Hidalgo, Adrian Cárdenas, Gabriel Quiroga
CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería)
Facultad Regional Mendoza, Universidad Tecnológica Nacional
Rodríguez 273, Mendoza, Argentina - 0261-5244576
(higiniofac,santiagocp)@frm.utn.edu.ar, (fabianhdlg,adriancard.mza)@gmail.com

Resumen

En los últimos años el tráfico de video ha crecido en forma exponencial, especialmente en los dispositivos móviles. Conocer el comportamiento del mismo y los requisitos necesarios de la red, ayudan a los administradores de la red a una mejor implementación. En este trabajo se propone un experimento de tráfico de video multicast en una red de laboratorio real como test bed. Se usa una topología en redes cableadas e inalámbricas con clientes inalámbricos trabajando en la norma IEEE 802.11ac. El tráfico de video se codificará en H264 y Theora, para evaluar y comprender el impacto sobre el tráfico de la red. Se analizará midiendo y comparando distintas métricas como cantidad de paquetes y bytes, espacios intertramas, tamaños de paquetes y tasa de bits efectiva para cada tipo de códec, etc. Además, se plantea un análisis estadístico de distribución de tramas. Este trabajo es una continuación de experimentaciones realizadas sobre redes cableadas para tráfico de video con codecs H261, H263 y H264. Las conclusiones del trabajo ayudarán a determinar las configuraciones a tener en cuenta para una adecuada gestión de redes similares y un uso eficiente de los recursos disponibles, sin comprometer el rendimiento y la QoS esperada..

Palabras clave: multicast, codecs, tráfico de video, Wi-Fi

Contexto

La línea de investigación está inserta en dos proyectos de análisis de tráfico multimedia, llevados adelante en el ámbito del Centro UTN CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería), del Departamento Ingeniería en Electrónica, de la Facultad Regional Mendoza, de la Universidad Tecnológica Nacional.

Los proyectos están dirigidos a investigar sobre: 1) Análisis de tráfico de video en redes cableadas e inalámbricas reales bajo las modalidades multicast y unicast, y 2) Análisis similares con simulación de tráfico de video en redes Wi-Fi bajo distintas normas 802.11 con el análisis correspondiente de QoS (calidad de servicio).

Introducción

El tráfico de aplicaciones de video en las redes se viene incrementando año tras año en forma permanente, ya sea en redes cableadas como inalámbricas. Indudablemente con el crecimiento exponencial del uso de dispositivos móviles, hay que tener en cuenta la disponibilidad del ancho de banda y recursos en las redes inalámbricas como Wi-Fi. Las mismas han mejorado su performance de acuerdo a las nuevas normas 802.11n/ac, etc; pero el incremento de tráfico y la cantidad de dispositivos, supera muchas veces las mejoras obtenidas. El uso de tráfico tipo multicast puede ser una solución interesante para una transmisión de datos simultánea a un grupo de usuarios, dado que ahorra recursos de la red,

enviando un único flujo de datos iguales a todos los receptores. Para el caso específico de tráfico de video, también es importante conocer el comportamiento de los distintos códecs; por lo cual se analizan 2 codecs ampliamente utilizados como son H264 y Theora.

Existen diferentes trabajos sobre los temas considerados, como priorización de tráfico de video en redes 802.11 [1], el análisis de tráfico multicast y protocolos de enrutamiento multicast [2], análisis de codecs de video [3], análisis de tráfico de video en diferentes contextos [4], análisis de tráfico de video conferencias multicast [5], un análisis comparativo de multicast vs unicast [6], el análisis de QoS para el tráfico de video [7 - 9], modelación de tráfico de video [10] y trabajar en IPTV multicasting [11]. Además incluimos trabajos de los autores en análisis de redes 802.11, tráfico de video y Calidad de Servicio [12-14]

Tráfico Multicast

El tráfico de red tradicional de unidifusión (o unicast) requiere que el emisor emita un flujo de datos individual a cada uno de los posibles receptores. A cambio, la multidifusión (o multicast) es una tecnología para conservar el ancho de banda, específicamente diseñada para reducir el tráfico, transmitiendo un único flujo de información potencialmente a miles de destinatarios (Figura. 1). De esta forma, se sustituyen las múltiples copias para todos los beneficiarios con la entrega de un único flujo de información. Por lo tanto, la multidifusión IP es capaz de reducir al mínimo la carga en el tráfico total de la red.

La multidifusión a través de redes inalámbricas es una función de comunicación fundamental, así como una meta desafiante. El objetivo es reducir sustancialmente el ancho de banda y la potencia consumida. Una red inalámbrica es por naturaleza una red de difusión o broadcast. Esto significa que un paquete puede ser interceptado por todos los nodos en el rango de transmisión del remitente. Por lo tanto, cada paquete se envía una sola vez y llegará a todos los destinatarios. Sin embargo,

a veces no todos los receptores están listos para recibir. Si el remitente tiene que esperar hasta que todos los receptores estén listos para recibir, entonces el sistema se vuelve inestable. Por otro lado, si el emisor transmite sin importar si los receptores están listos o no, puede producirse una pérdida grave de datos y de rendimiento. Multicast ofrece un equilibrio entre el rendimiento, la estabilidad y la pérdida de paquetes.

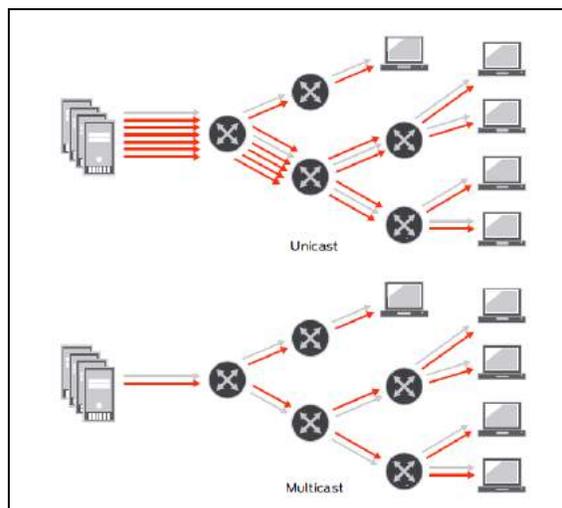


Figura 1: Comparación multicast vs unicast

Codecs

La compresión de video y audio es un facilitador fundamental para aplicaciones multimedia; y un aumento del número de algoritmos o códec estándares de la industria y de propiedad exclusiva están disponibles para que sea práctico almacenar y transmitir video en formato digital.

Los estándares de compresión están en continua evolución para hacer uso de avances en los algoritmos y tomar ventaja de los continuos aumentos en potencia de procesamiento disponible.

Para este trabajo se utilizarán 2 codecs:

- Theora es un códec de video abierto que está siendo desarrollado por la Fundación Xiph.org como parte de su proyecto Ogg (es un proyecto que pretende integrar el códec de video VP3 de On2, códec de audio Ogg

Vorbis y formatos de contenedores multimedia Ogg en una solución multimedia que puede competir con el formato MPEG-4. Theora se deriva directamente del códec VP3 de On2; actualmente los dos son casi idénticos, variando sólo en los encabezados de las tramas, pero Theora divergirá y mejorará el desarrollo principal de VP3 a medida que avanza el tiempo.

- b. H.264 o MPEG-4 parte 10 es una norma que define un códec de vídeo de alta compresión, desarrollada conjuntamente por el ITU-T Video Coding Experts Group (VCEG) y el ISO/IEC Moving Picture Experts Group (MPEG). Aunque H.264 utiliza las mismas técnicas de codificación generales que las normas anteriores, tiene muchas nuevas características que lo distinguen de las normas anteriores y se combinan para permitir una mayor eficiencia de codificación

Metodología

Los trabajos experimentales se realizarán sobre una red piloto de Laboratorio. La topología tendrá una composición mixta de redes cableadas e inalámbricas, routers, switches, access points, y equipos intermedios para dar soporte a los distintos tráficos. Además, equipos finales tanto cableados como inalámbricos que permitan simular una red empresarial de cierta envergadura.

El núcleo de la red tendrá como funcionalidad principal el ruteo de todos los tráficos, sobre el cual se realizarán las configuraciones de los distintos caminos del tráfico multicast, mientras que la conectividad primaria IP se realizará sobre cualquier protocolo de ruteo unicast para IPv6. En los bordes de esta red estarán los posibles usuarios que se podrán unir a los grupos multicast generados.

Se consideran distintas opciones de configuración, teniendo en cuenta la topología básica de la Figura 2. Para el caso de clientes

inalámbricos, los mismos se reemplazarán por los cableados, conectados a un Access Point correspondiente.

La inyección de tráfico de video multicast sobre la topología se realizará de diferentes maneras:

- Con un generador sintético como es el software IPTraffic, que permite generar distintas sesiones de multicast, sobre archivos de videos capturados.
- Con un emisor real, como una cámara de video IP que realice Streaming multicast y/o servidores de video multicast

Una vez realizadas las configuraciones básicas se generará tráfico de video multicast con distintos codecs y tipos de video; y se realizarán mediciones de rendimiento, cantidad de paquetes, jitter, errores, consumo de ancho de banda, retardo, etc.; para obtener los datos generales y sacar las conclusiones de performance buscadas.

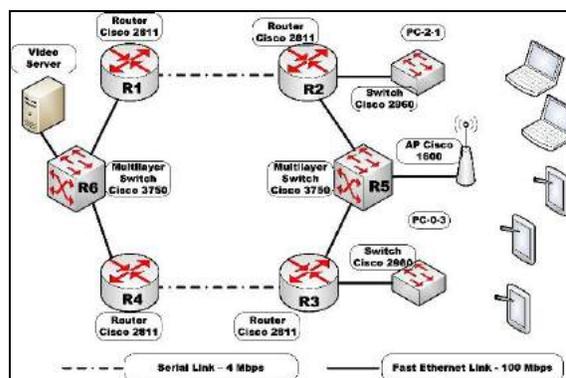


Figura 2: Topología de red

Líneas de Investigación y Desarrollo

Los temas y líneas de investigación, que se tratarán durante el desarrollo del proyecto son:

- Tráfico multicast.
- Direccionamiento multicast en IPv4/IPv6.
- Protocolos de ruteo multicast en IPV4/IPV6.
- Generación de streaming de video.
- Códec de video, especialmente en los actuales, como Theora, VP9, H.264, H.265, etc.

- Generación y análisis de métricas de resultados directas e indirectas.
- Análisis estadístico del tráfico de video.

Resultados Esperados

En proyectos anteriores de análisis de distintos tipos de tráfico, y específicamente de tráfico de video, se obtuvieron datos importantes en cuanto al funcionamiento y rendimiento de multicast frente a unicast en diferentes escenarios, contemplando redes cableadas e inalámbricas bajo el protocolo IPv4.

Para el tráfico de video se utilizaron archivos de video bajo distintos codecs, como MPEG4/2, H.264/3, obteniendo como datos primarios los siguientes:

- Cantidad de bytes y paquetes por códec,
- Tasa de bits,
- Tamaño de paquetes promedio,
- Espacio intertrama,
- Distribución estadística de paquetes por orden de llegada y de espacio intertrama, y
- Comportamiento frente a requisitos mínimos de QoS.

Siguiendo la línea se tiene en cuenta para el presente proyecto los siguientes objetivos:

- Analizar y comparar los distintos tipos de codecs de video actuales,
- Analizar y comparar las diferentes formas de generar tráfico de video multicast en IPv6
- Comparar el rendimiento del tráfico de video multicast frente a unicast en la medida que la cantidad de miembros multicast aumenta, variando las condiciones de la red

Obtenidos los datos primarios mencionados anteriormente, se buscarán los siguientes objetivos finales:

- Obtener conclusiones sobre la conveniencia de tráfico multicast para el transporte de video a múltiples usuarios dependiendo de tipos de tráfico de video, y

- Obtener conclusiones sobre el comportamiento y rendimiento de los distintos codecs de video de acuerdo a las distintas características de la red
- Obtener conclusiones sobre el rendimiento de los distintos tipos de tráfico de video según los codecs utilizados.

Formación de Recursos Humanos

El equipo de trabajo está integrado por docentes investigadores, y becarios graduados y alumnos del Centro UTN CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería) de la Facultad Regional Mendoza de la Universidad Tecnológica Nacional.

Se encuentra en desarrollo una tesis de Maestría. Además está la relación de la materia Proyecto Final de la carrera de Ingeniería en Electrónica, en la cual se incentiva que los proyectos finales de los alumnos estén enmarcados dentro de los proyectos de investigación y desarrollo del CeReCoN. Las actividades se llevan a cabo en el ámbito de las instalaciones del Centro, que cuenta con sus propias áreas de trabajo, 1 oficina técnico-administrativa, 2 Laboratorios con 11 computadoras cada uno, con material y con el siguiente equipamiento:

- 4 Routers CISCO 2811,
- 6 Routers CISCO 1721,
- 3 Switchs CISCO 2950,
- 2 Switchs CISCO 2960,
- 2 Switchs CISCO 3560,
- 1 ASA CISCO 5505,
- 2 routers Mikrotik,
- 4 Access Point Cisco y 2 Mikrotik,
- Placas inalámbricas de red,
- 2 cámaras de video IP con soporte de streaming multicast IPv4/IPv6,
- 1 Servidor de streaming de video
- 22 Computadoras con Sistemas Operativos Linux y Windows 7.
- Software IP Traffic de ZTI – Generador de tráfico IPv4/IPv6 unicast/multicast/broadcast y Medidor de performance (throughput,

cantidad de paquetes, jitter, número de errores, tanto enviados como recibidos).

- Hardware Air Pcap para captura de tráfico Wireless.
- Conexión a Internet por IPv4 e IPv6
- Servidor HP Proliant con Linux base y Máquinas Virtuales.
- Software VLC [15].
- Software Analizador Wireshark [16]

Referencias

- [1] Attilio Fiandrotti; Dario Gallucci; Enrico Masala; Enrico Magli "Traffic Prioritization of H.264/SVC Video over 802.11e Ad Hoc Wireless Networks" Computer Communications and Networks, 2008. ICCCN '08. Proceedings of 17th International Conference.
- [2] Ian Brown, Jon Crowcroft, Mark Handley, Brad Cain "Internet Multicast Tomorrow" - Articles of interest Vol 1 N°6 - <http://www.isoc.org/pubs/int/cisco-1-6.html>
- [3] A. Abdalla, A. Mazhar, M. Salah and Sahar Khalaf, "Comparative Study of Compression Techniques for Synthetic Videos," The International Journal of Multimedia & Its Applications (IJMA), vol. 6, no.2, April 2014.
- [4] D. Grois, D. Marpea, T. Nguyena and O. Hadarb, "Comparative Assessment of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC Encoders for Low-Delay Video Applications," SPIE Proceedings, vol. 9217, Applications of Digital Image Processing XXXVII, California, USA, September, 2014.
- [5] V. Chandrasekar, K. Baskaran – "Performance of Video Conferencing using Protocol Independent Multicast Routing with Core failure" - International Journal of Computer Applications (0975 – 8887) Volume 41– No.10, March 2012
- [6] Sorin Cocorada - Multicast vs. Unicast Transmissions for wireless IP Camera Surveillance Systems – Acta Technica Napocensis – Vol 48 N°3 2007
- [7] S. Tanwir, H. Perros and B. Anjum, "A QoS evaluation of video traffic models for H.264

AVC video," Fifth International Conference on Next Generation Networks and Services (NGNS), pp. 313-320, Casablanca, Morocco, May, 2014.

- [8] R. MacKenzie ; D. Hands ; T. O'Farrell "QoS of Video Delivered over 802.11e WLANs" Communications, 2009. ICC '09. IEEE International Conference.
- [9] Raj Kiran Addu and Vinod Kumar Potuvaradanam "Effect of Codec Performance on Video QoE for videos encoded with Xvid, H.264 and WebM/VP8" Master Thesis Electrical Engineering August 2014 Blekinge Institute of Technology Sweden
- [10] Aggelos Lazaris a, Polychronis Koutsakis "Modeling multiplexed traffic from H.264/AVC videoconference streams" Computer Communications 33 (2010) 1235–1242
- [11] Acosta Escobar M, Treminio Henriquez J, Estudio de IPTV multicasting para la Universidad de Don Bosco – Trabajo final Ingeniería en Ciencias de la Computación
- [12] S. Pérez, H. Facchini, A. Dantiacq, G. Cangemi and J. Campos, "An Evaluation of QoS for intensive video traffic over 802.11e WLANs-Quantitative analysis," International Conference on Electronics, Communications and Computers (CONIELECOMP), pp. 8-15, Cholula, Mexico, february, 2015.
- [13] S. Pérez, H. Facchini, A. Dantiacq, G. Cangemi and J. Campos, "Behavior of Codecs for Multicast Video Traffic using WAN Test Bed - Experimental Study," International Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON 2015), pp. 269-274, Santiago, Chile, October, 2015.
- [14] Santiago Pérez; Javier Campos; Higinio Facchini; Alejandro Dantiacq, " Experimental study of unicast and multicast video traffic using WAN test bed," 2016 IEEE Biennial Congress of Argentina (ARGENCON)
- [15] Software VLC, Available: <http://www.videolan.org/vlc/index.es.html>
- [16] Available: <https://www.wireshark.org/>.

SISTEMAS DE COMUNICACIÓN ROBUSTOS PARA INFRAESTRUCTURAS AVANZADAS DE MEDICIÓN DE ENERGÍA IMPLEMENTADAS SOBRE PLC (POWER LINE COMMUNICATIONS)

Jorge E. Veglia¹, David La Red Martinez¹, Reinaldo J. Scappini²

¹ Facultad de Ciencias Exactas, Naturales y Agrimensura, UNNE, Corrientes
{jeveglia, lrmdavid}@exa.unne.edu.ar

² Facultad Regional Resistencia, UTN, French 414 (3500) Resistencia, Chaco
rscappini@frre.utn.edu.ar

RESUMEN

Gran parte de las Infraestructuras Avanzadas de Medición (AMI, *Advanced Measurement Infrastructures*), desplegadas en los sistemas de distribución de energía eléctrica para el registro de los parámetros facturables y de calidad de servicio utilizan las mismas redes de energía como medio de comunicación (PLC, *Power Line Communications*). Para ello, utilizan la banda de frecuencias NB-PLC (*Narrowband PLC*, o PLC de banda estrecha, entre 90 y 500 KHz), con esquemas de modulación tales como la Multiplexación por División de Frecuencias Ortogonales (OFDM, *Orthogonal Frequency Division Multiplexing*). Dado que las redes de distribución de energía no se diseñan como canales de comunicación, presentan serios desafíos para la obtención de comunicaciones confiables. El proyecto procura alternativas que mejoren las prestaciones en cuanto a rendimiento y latencia en condiciones exigentes de atenuación y ruido en el canal. Se modela en Matlab/Simulink® el canal PLC de manera de incluir tales características en una configuración típica de distribución en Baja Tensión. Utilizando este modelo se analiza una propuesta tendiente a la mejora de sus prestaciones: la aplicación de una técnica de corrección de errores basada híbrida FEC-ARQ (Hybrid ARQ o HARQ) utilizada en los estándares de comunicaciones celulares inalámbricas. Mediante el mismo, se pretende analizar la variación de las métricas de rendimiento y latencia en distintos escenarios de las redes PLC.

Palabras clave: Redes inteligentes (*smart grids*); canales PLC; OFDM; ARQ Híbrido (HARQ)

CONTEXTO

Este proyecto se desarrolla como parte de la tesis académica de uno de los autores en el contexto de la Maestría en Sistemas y Redes de Telecomunicaciones, de la Facultad de Ciencias Exactas, Naturales y Agrimensura de la UNNE (Res. 797/12 C.S. U.N.N.E. y Res. 232/17 C.D. F.A.C.E.N.A. – U.N.N.E.)

1. INTRODUCCIÓN

Las infraestructuras avanzadas de medición constituyen un paradigma emergente en el campo de la medición de los servicios públicos de energía eléctrica, agua y gas. Una AMI está constituida por una red de comunicaciones de dos vías y es la integración de sensores avanzados, contadores inteligentes, sistemas de monitoreo, y sistemas de gestión de datos que permiten la recopilación y distribución de información entre medidores y utilidades (Güngör et al., 2011).

En términos generales, las AMI comparten una arquitectura básica, que puede sintetizarse en una red de área extensa (WAN, *Wide Area Network*), gobernada por un sistema central en el que residen las aplicaciones específicas de lectura, facturación, control, etc., a la que se vinculan un conjunto de redes de área local (LAN, *Local Area Network*). En las soluciones actuales, estas LAN utilizan dos medios físicos principales para la

comunicación dentro de la red inteligente: radiofrecuencia (RF), mediante la utilización de la banda libre ISM (*Industrial, Scientific and Medical*, 902-928 MHz; 2,400-2,483 GHz); y la misma red eléctrica, mediante la llamada comunicación por líneas de potencia o PLC. En la actualidad existen desarrollos comerciales que hacen uso de uno u otro; sin embargo, desde el punto de vista de las empresas eléctricas, la utilización de la red de distribución de energía resulta muy atractiva, dado que además de su ubicuidad, es un medio de su propiedad, lo que evita la necesidad de tendidos adicionales y sobrecostos de comunicación.

Obviamente, la red eléctrica no se diseña como una red de comunicación, y esto significa que muchos de sus parámetros característicos y sus detalles constructivos distan mucho de ser ideales para transmitir información. Los problemas de atenuación, desadaptación de impedancias y ruido de variadas fuentes, tipos e intensidades limitan de manera significativa la capacidad de la red. A pesar de ello, y dado que el tráfico de datos es esporádico y de un ancho de banda muy reducido, con velocidades tan bajas como unos pocos cientos o miles de kilobits por segundo, las funciones básicas de las AMI pueden implementarse en una LAN típica de pocos cientos de medidores.

Sin embargo, el desarrollo de la electrónica utilizada en los medidores hace que la cantidad de información que pueden registrar y almacenar se incremente cada vez más, con sobrecostos mínimos. Como consecuencia, las normas regulatorias van exigiendo que esta información (por ejemplo, la curva de consumo) sea puesta a disposición de los clientes por parte de las empresas prestatarias, de manera que estos puedan tele gestionar sus demandas. Al mismo tiempo, las exigencias de calidad de servicio, así como el creciente desarrollo de las tecnologías “verdes” como la generación solar domiciliaria y el intercambio bidireccional de energía con los clientes obligan a las empresas a instalar en la red

eléctrica otros dispositivos que aportan importantes volúmenes de datos a la red. Para adaptarse a los nuevos escenarios, y tomando como referencia el modelo de capas del estándar básico de referencia de Interconexión de Sistemas Abiertos (OSI), es necesario lograr mejores desempeños a nivel de capa física (PHY) y de la capa de control de acceso al medio (MAC) (Sendin et al., 2014).

Los estándares actuales de PLC (por ejemplo G3 y PRIME, y el más reciente ITU-T 1901.2) avanzan en esta dirección al incorporar técnicas avanzadas de modulación multiportadora, como OFDM. En todos ellos, para afrontar la ocurrencia de errores en la transmisión debidas a las características no ideales del canal, se utilizan técnicas de manipulación de la información binaria que tienden a mejorar la relación señal/ruido (Eb/No). Si esto no resulta suficiente, acuden a métodos de corrección de error hacia adelante (FEC, *Forward Error Correction*). Su performance ha sido testeada en distintas simulaciones y pruebas de campo en redes estándar con resultados generalmente aceptables. Sin embargo, no se ha probado si las soluciones comerciales actuales están adecuadamente preparadas para infraestructuras físicas más “duras”, como lo pueden resultar redes eléctricas más precarias o en ambientes rurales o suburbanos. En términos generales, y dado que los canales PLC son ruidosos y cambiantes, una opción conservadora de modulación de menor riesgo y baja velocidad es apropiada para un despliegue sobre el terreno. Sin duda es funcional para las reducidas prestaciones de las AMI actuales en redes de un estado razonablemente adecuado, pero se puede suponer que resultarán insuficientes para escenarios más exigentes, tanto de tráfico como de condiciones desfavorables de la red. Al menos una de los estándares (PRIME) hace uso también del método de repetición automática (ARQ, *Automatic Repeat-reQuest*) para mejorar las tasas de recepción de paquetes. Sin embargo, la investigación y los

restantes estándares en PLC no parecen haber dirigido su atención en este sentido, y en particular a una técnica de corrección de errores que sí ha ganado amplio desarrollo en el campo de la RF, como lo es la combinación de FEC con ARQ, conocida como ARQ híbrida (HARQ). Esta técnica forma parte de los estándares inalámbricos de banda ancha actuales tales como IEEE 802.16, utilizados en 3GPP LTE (*Long Term Evolution*) y WiMAX, por caso. Los interesantes resultados reportados en la bibliografía en cuanto a mejoras en el rendimiento y la latencia en canales que presentan desafíos similares (ruido, multitrayecto, etc.), -si bien poseen una naturaleza física diferente y operan en un rango de frecuencia distinto-, llevan a la idea que su aplicación puede resultar provechosa en los canales PLC, con vistas a mantener o mejorar las prestaciones de la red en condiciones más desfavorables del canal.

Por lo tanto, y con la motivación de la búsqueda de técnicas de comunicación que proporcionen un mayor rendimiento y robustez tal que permitan superar las dificultades planteadas por las características propias de los canales PLC y sus variadas condiciones de entorno, el proyecto pretende desarrollar un modelo de simulación en la capa física con una modulación OFDM en el espectro NB-PLC, con una técnica de corrección de errores basada en la filosofía HARQ, orientado a su utilización en las AMI.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto requirió el desarrollo de un modelo sobre el que pueda ser evaluada la técnica de corrección de errores propuesta. Para ello, fue necesario elegir en primer lugar un modelo de canal PLC. Las líneas de investigación en este sentido hacen uso de dos enfoques principales: una aproximación top-down o descendente y una bottom-up o ascendente. Siguiendo el resumen de W. Zhu

(2014), los enfoques top-down, se basan en un método empírico, en el que se toman medidas en una red real con las que se ajustan diferentes parámetros que constituyen la función de transferencia. Es un método eficiente, aunque afectado por los errores de medida. Su ventaja es que resulta un modelo de baja complejidad. La desventaja más significativa es su baja flexibilidad. El modelo obtenido para una red específica o una banda de frecuencia no pueden aplicarse a otras redes y bandas de frecuencia, al tiempo que carecen de conexión física con la realidad.

La segunda opción, los enfoques bottom-up, se basan en la teoría de líneas de transmisión y consisten en un estudio analítico de la red y sus componentes. Conocidos la topología, tipos de cables y cargas de la red, se modelan matemáticamente para generar un modelo discreto de canal, que tiene en cuenta todo tipo de retrasos, atenuaciones y reflexiones. La ventaja de este enfoque es que puede aplicarse a diversas situaciones, siempre que la información de red esté disponible. Sin embargo, también tiene desventajas: su complejidad computacional crece con el tamaño de la red y la recopilación de los elementos de red antes mencionados es difícil. Sin embargo, y como el mismo Zhu menciona, el problema más importante del modelado de canales para las comunicaciones de redes inteligentes es la falta de modelos de canal de banda estrecha. La mayoría de los modelos de canal, especialmente los top-down, son modelos de canal de banda ancha en la banda de frecuencias de 1-100MHz. Muy pocas de ellas se centran o cubren la NB-PLC (3-500 kHz).

Por lo tanto, y en gran parte con la motivación de obtener una herramienta que permita modelar las características del canal en función de la topología de los componentes de la red eléctrica y las características de sus cargas sin un alto costo computacional ni de desarrollo, se optó por una tercera opción, la de implementar un modelo de red utilizando la librería Simscape de Simulink. Gracias a

ello puede ser modificado de acuerdo a las topología y características de los componentes de la red con mucha facilidad. Siguiendo las ideas de trabajos como Masood et al (2016), se implementó el modelo ilustrado en la Figura 1. Las simulaciones de su comportamiento en cuanto a respuesta en frecuencia y atenuación resultan en general concordantes con los (pocos) resultados disponibles obtenidos con los otros dos métodos. A este modelo se le añade otro que incorpora las dos fuentes principales de ruido eléctrico presentes en los canales PLC: el ruido aditivo blanco gaussiano (AWGN, *Additive White Gaussian Noise*) y el ruido impulsivo asincrónico. En cada uno de estos módulos, es posible configurar la relación señal/ruido y la intensidad de los impulsos respectivamente, de manera de poder configurar distintos escenarios. Finalmente, fue necesario modelar el sistema transmisor/receptor, para el que se utilizaron las principales especificaciones del estándar PRIME. El sistema completo según se describió se representa en la Figura 2.

En razón de que el principal problema a resolver en todo sistema de comunicación es asegurar la confiabilidad de la información transmitida por el canal, de modo de asegurar una integridad razonable dada por una tasa de error predeterminada, se analizarán los principales métodos de corrección de errores utilizados. En particular, el proyecto propone la utilización de una técnica que sigue las ideas fundamentales de técnica conocida como ARQ híbrida (HARQ), que combina la solicitud de repetición automática (ARQ) con la corrección de error hacia adelante (FEC). El desarrollo implica la programación en Matlab de los algoritmos necesarios de manera de implementar la técnica de corrección de errores sobre el modelo descrito de canal y ruido asociado. Se elaborarán las métricas que permitan comparar el desempeño de la mejora propuesta en términos de rendimiento, utilización del ancho de banda y tasas de error, para diferentes configuraciones de la topología de la red y para distintas condiciones de ruido en el canal.

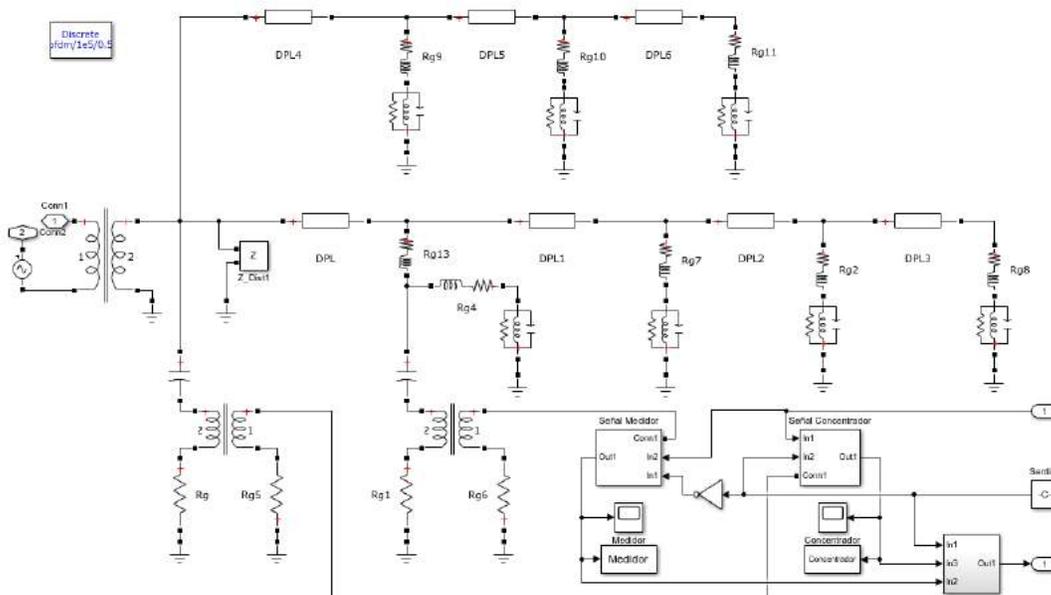


Figura 1. Modelo de canal PLC y módems en recepción y transmisión

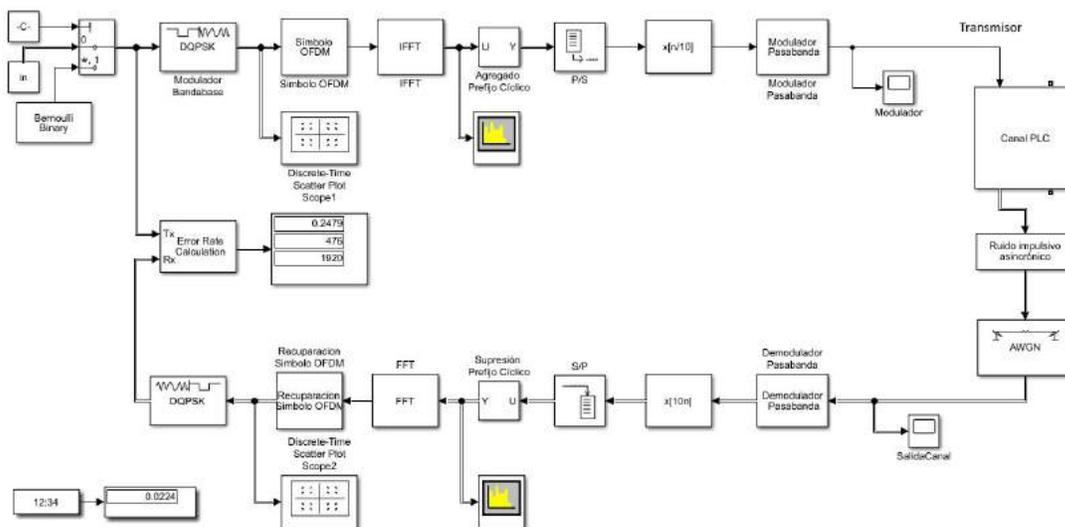


Figura 2. Esquema básico Transmisor/Receptor OFDM con el canal y las fuentes de ruido

3. RESULTADOS OBTENIDOS / ESPERADOS

El proyecto pretende arribar a conclusiones respecto a la aplicabilidad de la técnica de comunicación propuesta, analizando las tasas de error obtenidas y otras métricas útiles en función de las características particulares del canal físico. Al mismo tiempo, se espera producir una herramienta de simulación que pueda ser optimizada en función de resultados experimentales o trabajos de investigación que propongan mejores modelos del canal, y otros esquemas de codificación o de corrección de errores que permitan obtener mejores resultados.

Adicionalmente, se espera producir un material de referencia para contenidos curriculares o nuevas investigaciones.

4. FORMACIÓN DE RECURSOS HUMANOS

Se espera las líneas de trabajo que surjan a partir del proyecto promuevan la formación de grupos de investigación y desarrollo sobre tales tópicos entre los docentes y alumnos de las carreras de Ingeniería en Electrónica y la Licenciatura en Sistemas de Información de la FA.C.E.N.A.

5. BIBLIOGRAFÍA

- Cheng, J., "Coding Performance of Hybrid ARQ Schemes", IEEE Transactions on Communications, Vol. 54, N° 6, pp. 1017-1029, 2006
- Gungor, V.C.; Sahin, D.; Kocak, T.; Ergut, S.; Buccella, C.; Cecati, C.; Hancke, G.P., "Smart Grid Technologies: Communication Technologies and Standards", IEEE Transactions of Industrial Informatics, Vol. 7, N° 4 Nov. 2011, pp. 529-539, 2011
- Masdod, B; Haider, A; Baig, S., "Modeling and Characterization of low Voltage Access Network for Narrowband Powerline Communications", Journal of Electric Engineering Technologies (2106), en <http://dx.doi.org/10.5370/JEET.2016.11.192>
- Sendin, A.; Peña, I.; Angueira, P., "Strategies for Power Line Communications Smart Metering Network Deployment", Energies 2014, 7, 2377-2420, doi: 10.3390/en7042377, disponible en <http://www.mdpi.com/1996-1073/7/4/2377>, 2014
- Zhu, W; "Power Line Communications over Time-Varying Frequency-Selective Power Line Channels for Smart Home Applications", (2014), https://livrepository.liverpool.ac.uk/..../1/ZhuWen_July2014.pdf

Nuevo Modelo de Decisión para Gestión de Tráfico en Redes

Domingo A. Rios¹, David L. la Red Martínez²

¹Facultad de Ciencias Exactas, Naturales y de Agrimensura
Av. Libertad 5470, (3400) Corrientes, Argentina, +54-379-224493
domingoalbertorios@hotmail.com

²Facultad de Ciencias Exactas, Naturales y de Agrimensura
Av. Libertad 5470, (3400) Corrientes, Argentina, +54-379-4638194
lrmdavid@exa.unne.edu.ar

Resumen

En el tráfico de redes es necesario que los nodos deban tomar decisiones basados en acuerdos respecto del acceso a rutas disponibles; las decisiones pueden estar relacionadas con el estado de los nodos de acuerdo a, por ejemplo el porcentaje de CPU usado, la memoria disponible y el número de paquetes encolados para ser distribuidos, como así también el estado de los tramos entre nodos.

Así surge la siguiente pregunta: ¿cuáles son los modelos de decisión y los operadores de agregación que habrá que generar para la toma de decisiones en la gestión de tráfico en redes de paquetes de datos, que trasciendan el enfoque tradicional de la gestión de tráfico en redes, teniendo en cuenta la auto-regulación?

Los modelos considerarán la posibilidad de imputación de datos faltantes y la fuzzyficación de ciertas variables, utilizando operadores OWA, buscando generar operadores de agregación específicos.

Los modelos desarrollados se evaluarán comparando sus características con los modelos habitualmente utilizados.

Palabras clave: Comunicaciones de Datos; Transmisión de Paquetes de Datos; Modelos de Decisión; Operadores de Agregación.

Contexto

Se considera especialmente importante estudiar la aplicación de modelos de

decisión para la toma de decisiones en grupo que se desprendan de conceptos de sistemas complejos auto-regulados.

Se pretende generar nuevos modelos de toma de decisiones en el tráfico de paquetes, contemplando además la aplicación de métodos de imputación de datos para aquellos casos de datos faltantes, por ejemplo, como consecuencia de problemas en las comunicaciones entre los nodos, y fuzzyficación de variables para dar soporte a situaciones donde no es posible o conveniente expresar valores exactos.

Una situación de congestión se entiende como la consecuencia de una degradación de las prestaciones de la red que lleva a una situación de colapso provocado por el incremento progresivo de tráfico de retransmisiones de paquetes perdidos, (Floyd and Fall, 1999), (Jacobson, 1988). Este hecho puede llevar a que el porcentaje de paquetes que llegan a su destino disminuya dramáticamente conforme crece el tráfico generado desde los nodos extremos de la red.

El problema de la gestión del tráfico y de pertinente control de congestión puede enfocarse matemáticamente desde el punto de vista de la teoría de control de procesos.

En esta propuesta se tiene previsto desarrollar modelos de decisión de ruteo que de manera dinámica y colaborativa detecten y evadan zonas congestionadas de la red de datos.

Introducción

Los nuevos servicios y tecnologías surgidos durante los últimos años en Internet plantean serios problemas para los

mecanismos de control de congestión y gestión del ancho de banda y otros recursos de la red desarrollados hasta la fecha, (Floyd, 2000). Es más, el escaso conocimiento que se tiene sobre la dinámica del tráfico de Internet, así como la falta de enfoques de análisis y técnicas de ingeniería del tráfico en redes IP, suponen una importante limitación para la estabilidad y seguridad global de la red, el despliegue de nuevos servicios y el uso eficiente de nuevas tecnologías de comunicaciones.

Actualmente, existe un convencimiento generalizado acerca de la necesidad de modelos y herramientas con mayor base teórica y experimental que las disponibles para el estudio de la dinámica del tráfico en Internet, (Aiken et al, 2002) y (Floyd and Kohler, 2002).

Teniendo en cuenta las necesidades y los avances producidos en una sociedad sumamente compleja, resulta de gran importancia destacar tanto la transmisión de información, como la necesidad de que ésta llegue a destino en el momento preciso mediante el uso de las redes. De hecho, todas las sociedades, por definición, han sido y serán “sociedades de la comunicación”.

La capa de red, dentro de una arquitectura de red de datos, es la que se encarga de llevar los paquetes de datos desde el origen (estación transmisora) hasta el destino (estación receptora). El algoritmo de ruteo utilizado debe cumplir con ciertas propiedades que aseguren su eficiencia: corrección, estabilidad, robustez, equitatividad, sencillez y optimalidad.

Algunos protocolos de enrutamiento mantienen tablas de enrutamiento dinámicas por medio de mensajes de actualización de las rutas, que contienen información acerca de los cambios sufridos en la red, y que indican al software del router que actualice la tabla de enrutamiento en consecuencia. Intentar utilizar el enrutamiento dinámico sobre situaciones que no lo requieren es una pérdida de ancho de banda, esfuerzo, y en consecuencia de dinero.

La proliferación de las redes informáticas, en las cuales existe mucho tráfico de paquetes de datos, hace necesario disponer de modelos de decisión que permitan a los paquetes ser ruteados de

acuerdo a su prioridad y a los criterios de optimización adoptados, para ser enviados por la mejor ruta disponible para llegar a destino. Para tomar decisiones son necesarios diferentes niveles de acuerdo entre los nodos involucrados en las distintas rutas posibles, considerando, por ejemplo, el estado actual del nodo, esto incluye el porcentaje utilizado de CPU, porcentaje de memoria usada, número de paquetes encolados, un indicador de saturación de los tramos de los enlaces, etc.

El presente proyecto pretende generar nuevos modelos de toma de decisiones en el tráfico de paquetes, contemplando además la aplicación de métodos de imputación de datos para aquellos casos de datos faltantes, por ejemplo, como consecuencia de problemas en las comunicaciones entre los nodos, y fuzzyficación de variables para dar soporte a situaciones donde no es posible o conveniente expresar valores exactos.

Líneas de Investigación, Desarrollo e Innovación

Se han realizado y publicado varios trabajos relacionados con las áreas del conocimiento comprendidas en el presente proyecto; dichos trabajos se agrupan según se indica a continuación:

Comunicación en sistemas distribuidos

En (La Red Martínez y Goitia, 2004) se describen los principales algoritmos de comunicación en sistemas distribuidos (algoritmos clásicos de las ciencias de la computación).

En (La Red Martínez y Agostini, 2014) se presenta un entorno de enseñanza – aprendizaje de las comunicaciones de datos mediante animaciones.

Métodos de imputación de datos

En (Peláez et al., 2008) se describen métodos de imputación generales y para aplicaciones específicas.

Operadores de agregación

En (Doña et al., 2011) se presenta un modelo de decisión en grupo con la utilización de operadores de agregación de la familia OWA (Ordered Weighted Averaging: Promedio de Pesos Ordenados).

En (La Red Martínez y Acosta, 2015) se presentan las principales propiedades matemáticas y las medidas de

comportamiento relacionadas con los operadores de agregación.

En (La Red Martínez y Pinto, 2015) se presenta una revisión acerca de los operadores de agregación, especialmente los de la familia OWA.

En (Pelález et al., 2003, 2004, 2009) se analizan operadores de agregación (de mayoría) en grupo que buscan la representación de la mayoría.

Toma de decisiones en grupo

En (La Red et al., 2011a) se presenta el operador WKC-OWA para agregar información en problemas de decisión democrática.

En (La Red et al., 2011b) se presenta un modelo de decisión en grupo con la utilización de etiquetas lingüísticas y una nueva forma de expresión de las preferencias de los decisores.

En (La Red Martínez y Acosta, 2014) y en (La Red Martínez et al., 2014) se presenta una nueva perspectiva para los modelos de decisión para la sincronización de procesos en sistemas distribuidos.

En (La Red Martínez y Acosta, 2015) se realiza una revisión del modelado de preferencias para los modelos de decisión.

Resultados y Objetivos

Resultados

El presente proyecto intenta desarrollar nuevos modelos de decisión y operadores de agregación para la gestión de tráfico en redes para escenarios dinámicos (cargas de tráfico variables) contemplando cambios en la topología (agregado y caída de nodos y enlaces) para los siguientes escenarios:

** Que los paquetes de datos sean despachados considerando como estado de cada ruta el promedio de los estados de sus nodos.

** Que los paquetes de datos sean despachados considerando como estado de cada ruta el estado de su nodo y enlace más cargado.

** Que los paquetes de datos sean despachados considerando como estado de cada ruta al promedio de los estados de un cierto porcentaje de sus nodos y enlaces más cargados.

La investigación que se realizará será de tipo teórica en la etapa de desarrollo de los modelos de decisión. Una vez definidos

teóricamente los modelos de decisión antes mencionados, se procederá a la validación de estos comparando sus prestaciones con las de los modelos de enrutamiento habitualmente utilizado en las redes de datos.

El sistema de matrices de datos que se utilizará contemplará las siguientes premisas y estructuras de datos. Se trata de asignar rutas a paquetes de datos, debiendo decidirse cual ruta va a ser la mejor, en base a un criterio establecido, del estado de los nodos y enlaces.

Se tendrá en cuenta:

Conjunto de nodos $1, \dots, n$.

Conjunto de enlaces (tramos) $1, \dots, e$.

Conjunto de rutas $1, \dots, r$, constituidas por un subconjunto de nodos y un subconjunto de enlaces.

Estado posible de cada uno de los n nodos:

- Número de paquetes encolados en el nodo.
- Prioridades de los paquetes.
- Uso de CPU.
- Uso de memoria.
- Tamaño de paquetes (número t de paquetes).
- Predisposición y decisión (prioridad global) para otorgar el acceso a cada una de las r rutas.

Estado posible de los e enlaces:

- Capacidad utilizada.
- Valor-Peso.

Estado posible de las r rutas:

- Se calculará con los operadores de agregación en base al estado de los nodos y enlaces constituidos.

Se efectuarán simulaciones con los modelos de decisión propuestos, a los efectos de analizar el comportamiento de estos para las mismas condiciones de carga de tráfico.

Se considerará la situación de datos faltantes, y la posibilidad de utilizar métodos de imputación de datos para completar los datos faltantes.

Objetivo general del proyecto

Obtener el conocimiento necesario para desarrollar modelos de decisión aplicables a la gestión de tráfico en redes.

Objetivos específicos del proyecto

- Se pretende generar nuevos modelos de toma de decisiones en el tráfico de

paquetes, contemplando además la aplicación de métodos de imputación de datos para aquellos casos de datos faltantes, por ejemplo, como consecuencia de problemas en las comunicaciones entre los nodos, y fuzzyficación de variables para dar soporte a situaciones donde no es posible o conveniente expresar valores exactos.

Formación de Recursos Humanos

El equipo de trabajo está integrado por un Doctor y un Licenciado en Sistemas de Información.

Actualmente se está trabajando en la elaboración de la tesis de la Maestría de Sistemas de Redes y Telecomunicaciones siendo el tema “Nuevo Modelo de Decisión para Gestión de Tráfico en Redes” cuyo plan de tesis fue aprobado por el Consejo Directivo de la Facultad de Ciencias Exactas y Naturales y Agrimensura de la UNNE, mediante Resolución N°0640/17 de fecha 10 de agosto de 2017.

Referencias

- Aiken B., Bahl V., Bhattacharjee B., Braden B., et al. (2002). Report of NSF Work-shop on Network Research Testbeds. Technical report, National Science Foundation, Directorate for Computer and Information Science and Engineering, Advanced Networking Infrastructure & Research Division. [Online]. Disponible: http://gaia.cs.umass.edu/testbed_workshop.
- Doña, J. M.; Gil, A. M.; Peláez, J. I.; La Red Martínez, D. L. (2011). A System Based on the Concept of Linguistic Majority for the Companies Valuation. *Revista Econo Quantum*. V. 8 N° 2. Pp. 121-142. México.
- Floyd S. (2000). Congestion Control Principles. RFC2914, Network Working Group, Best Current Practice.
- Floyd S. y Kohler E. (2002). Internet Research Needs Better Models. In *ACM SIGCOMM First Workshop on Hot Topics in Networks (HotNetsI)*. Princeton, New Jersey. USA. [Online] Disponible :<http://www.icir.org/floyd/papers.html>.
- Floyd S. and Fall K. (1999). Promoting the Use of End-to-End Congestion Control in the Internet. *IEEE/ACM Transactions on Networking*, 7(4):458–472. [Online]. Disponible: <http://www.icir.org/floyd/papers.html>.
- Jacobson V. (1988). Congestion Avoidance and Control. *ACM Computer Communication Review SIGCOMM '88 Symposium: Communications Architectures and Protocols*, 18(4):314–329. [Online]. Disponible: <http://citeseer.ist.psu.edu/jacobson88congestion.html>.
- La Red Martínez, D. L., Acosta, C. (2015). Aggregation Operators Review - Mathematical Properties and Behavioral Measures; Volume 7 – N° 10; *International Journal of Intelligent Systems and Applications (IJISA)*; pp. 63-76; Hong Kong.
- La Red Martínez, D. L., Acosta, J. C. (2014). Perspectives of New Decision Making Models of Processes Synchronization in Distributed Systems, Volume 09 – N° 1, *International Journal of Management and Information Technologies*, pp. 1504-1512, ISSN N° 2278-5612, U.S.A.
- La Red Martínez, D. L., Acosta, J. C. (2015). Review of Modeling Preferences for Decision Models, Volume 11 – N° 36, *European Scientific Journal (ESJ)*, pp. 1-18, ISSN N° 1857-7881, Macedonia.
- La Red Martínez, D. L., Agostini, F. (2014). ISO/OSI model and data communication by animations, *Revista WIT Transactions on Information and Communication Technologies*, WIT Press, Vol. 58, pp 963-970, ISBN N° 978-1-84564-975-3, Ashurst, Southampton, U.K.
- La Red Martínez, D. L., Goitia, M. J. (2004). Protocolo de Enrutamiento. *Simulador de Trafico de Redes*. V Jornadas de Informática y Sociedad; Libro de Actas, pag. 425-432; ISBN N° 84-7485-927-1; España.
- La Red Martínez, D. L., Pinto, N. (2015). Brief Review of Aggregation Operators;

Volume 22 – N° 4; Wulfenia Journal; pp. 114-137; Austria.

• La Red, D. L.; Doña, J. M.; Peláez, J. I.; Fernández, E. B. (2011a). WKC-OWA, a New Neat-OWA Operator to Aggregate Information in Democratic Decision Problems. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, World Scientific Publishing Company, págs. 759-779. Francia.

• La Red, D. L.; Peláez, J. I.; Doña, J. M. (2011b). A Decision Model to the Representative Democracy With Expanded Vote. V. 1 N° 1. *Revista Pioneer Journal of Computer Science and Engineering Technology*. Pp. 35-45. India.

• Peláez, J.I., Doña, J.M., La Red Martínez, D.L. (2003). Analysis of the Majority Process in Group Decision Making Process. 9th International Conference on Fuzzy Theory and Technology. North Carolina. USA.

• Peláez, J.I., Doña, J.M., La Red Martínez, D.L. (2008). Fuzzy Imputation Method For Database Systems. En Galindo, J. (Ed.). *Handbook of Research on Fuzzy Information Processing in Database*. Hershey. Information Science Reference. USA.

• Peláez, J.I., Doña, J.M., La Red Martínez, D.L. (2009). A Mix Model of Discounted Cash-Flow and OWA Operators for Strategic Valuation. V. 1 N° 2. *Revista Interactive Multimedia and Artificial Intelligence - Special Issue On Business Intelligence And Semantic Web*. Pp. 20-25. España.

• Peláez, J.I., Doña, J.M., La Red Martínez, D.L., Mesas, A. (2004). Majority Opinion in Group Decision Making Using the QMA-OWA Operator. *Proceeding of ESTYLF*. 449-454.

Metodología para la selección de recursos computacionales gestionados con tecnologías de cloud computing en ambientes educativos

Hugo R. Haurech¹, David L. la Red Martínez²

¹Facultad de Ciencias Económicas / Universidad Nacional de Misiones
Ruta 12 km 7.5, (3300) Posadas, Misiones, Argentina,
haurech@fce.unam.edu.ar

²Facultad de Ciencias Exactas, Naturales y de Agrimensura / Universidad Nacional del Nordeste
Av. Libertad 5470, (3400) Corrientes, Argentina,
lrm david@exa.unne.edu.ar

Resumen

Debido a los avances tecnológicos, las organizaciones educativas se enfrentan a diversos retos al proporcionar el apoyo mediados por las TI para el desarrollo de las actividades de índole académicas y administrativas. Ante el desafío que representa adoptar tecnologías nuevas y emergentes, resulta imperioso el uso eficaz de ellas en el proceso de enseñanza y aprendizaje, a fin de brindar una educación acorde a los requerimientos actuales. Se deben "explotar las oportunidades que ofrece el cloud computing (CC)" (V. H. Pardeshi, 2014), para permitir el acceso conveniente a un conjunto compartido de recursos computacionales configurables (por ejemplo, redes, servidores, almacenamiento, aplicaciones y servicios) que se pueden aprovisionar y liberar rápidamente con un esfuerzo mínimo de gestión o una interacción con el proveedor de servicios.

De forma similar y con mayor precisión el National Institute of Standards and Technology (NIST), lo define como "un modelo tecnológico que permite el acceso ubicuo, adaptado y bajo demanda en red a un conjunto compartido de recursos de computación configurables -redes, servidores, equipos de almacenamiento, aplicaciones y servicios-, que pueden ser rápidamente aprovisionados y liberados con un esfuerzo de gestión reducido o interac-

ción mínima con el proveedor del servicio" (P. Mell and T. Grance, 2011).

De acuerdo a ello surgen interrogantes sobre, la elección del modelo adecuado para el despliegue del entorno y qué servicio se ajusta a las necesidades del ambiente, las tecnologías vigentes y la seguridad e integridad de la información.

Palabras clave: Cloud computing – Recursos computacionales – Redes computacionales – Entornos educativos.

Contexto

Este estudio está relacionado con el trabajo final "Metodología para la selección de recursos computacionales gestionados con tecnologías de cloud computing en ambientes educativos" de la Maestría en Tecnologías de la Información, el cual se desarrollará en la Facultad de Ciencias Económicas, siendo la Dirección de Tecnologías de la Información el área donde se concentrará los mayores esfuerzos, involucrando a sectores específicos de la unidad académica, como son las cátedras relacionadas con informática, que serán las principales fuentes de consulta sobre los requerimientos a cubrir con una solución de CC generable a partir de la aplicación de la metodología planteada, para delinear los requerimientos.

Una tarea común entre los integrantes de una unidad académica es compartir archivos para lo cual suelen utilizarse métodos que involucran medios extraíbles o archivos adjuntos a un correo electrónico lo cual posee como limitante, en principio la capacidad que se requiere, del medio y del gestor de correos para incorporar archivos adjuntos, con el consecuente aumento en volumen del contenido de los agentes de transferencia correo (MTA: Mail Transfer Agent). Además esa información que puede ser de interés general queda confinada entre los participantes por la carencia de una adecuada gestión del conocimiento.

Los archivos que se desean compartir pueden clasificarse de acuerdo a las tareas que desarrolla e intercambia el personal de los distintos claustros, contenidos que son particulares y afines de cada área, y otros que poseen formatos estandarizados como planillas, notas, solicitudes y encuestas pre-elaboradas que son de uso común entre los miembros de la unidad académica y por tal motivo pueden ser adquiribles de manera simple y sin necesidad de mediar con algún interlocutor.

Esta situación también se da en el plano académico, más específicamente en el ámbito de enseñanza y aprendizaje, con material para el desarrollo y presentación de trabajos prácticos, tesis o guías de estudio con formato definido. A esto se suman las diferentes estrategias docentes que incluyen la organización de sus cátedras con material audiovisual que requiere del acceso a Internet sin filtros y contar con la capacidad relacionado al ancho de banda, con la consecuente exigencia de infraestructura y configuraciones que demandan tareas de aspecto técnico. Las actividades planteadas por las cátedras para desarrollarse de manera colaborativa deben llevarse a cabo en ambientes externos a la organización, siendo el proveedor del servicio el encargado de brindar el soporte en cuanto a los recursos.

En cuanto a la edición de documentos, sean de texto o planillas de cálculo, se pudo evidenciar el uso de diferentes plataformas ofimáticas (Office de Microsoft, LibreOffice, OpenOffice) y de versiones, lo cual implica

que el intercambio de archivo pueda estar sujeto a modificaciones de formato seguido de errores de interpretación y calificación.

Introducción

En la actualidad existen distintos proveedores que brindan servicios en la “nube”, también aplicaciones a partir de las cuales se pueden acceder a herramientas para el trabajo colaborativo. En ambos casos el denominador común es el modelo de CC, el cual está dividido en dos grandes grupos: Modelo de servicio y Modelo de despliegue, el primero de ellos hace referencia a servicios que pueden contener y ser accedidos, caracterizados como SaaS (Software as a Service), PaaS (Platform as a Service) y IaaS (Infrastructure as a Service).

El modelo SaaS es el servicio de más alto nivel de abstracción para el usuario – cliente, se trata de la provisión de aplicaciones que se encuentran desplegadas sobre una infraestructura de CC. El usuario accede a las mismas a través de un navegador web, sin tener ninguna posibilidad de administración de la infraestructura subyacente ni sobre elementos avanzados de la configuración de las aplicaciones utilizadas. El modelo IaaS consiste en la provisión de servicios de infraestructura TI (Tecnologías de la Información) como las máquinas virtuales, los recursos de red, el espacio de almacenamiento, la capacidad de procesamiento y otras soluciones que tienen una marcada orientación al sector de administración de TI. Este modelo entrega los elementos necesarios para la ejecución de aplicaciones, incluyendo en ocasiones la provisión de un sistema operativo (SO), además de algún producto software específico para la gestión del conjunto de recursos demandados por el cliente. El modelo PaaS, consiste en un modelo en el que se provee un conjunto de herramientas de software orientadas a desarrolladores de software. Permite el despliegue de aplicaciones en una infraestructura de CC de forma transparente, abstrayendo al desarrollador de la gestión de la misma. Dependiendo del proveedor, serán accesibles diver-

los lenguajes de programación y entornos o herramientas de desarrollo.

El segundo modelo se refiere a la localización y gestión del cloud como plataforma dentro de una infraestructura capaz de soportar servicios, se divide en cuatro grupos según el ámbito a ser desplegados como: Pública, Privada, Híbrida y Comunitaria (P. Mell and T. Grance, 2011), (S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, and A. Ghalsasi, 2011).

La infraestructura Pública se trata de una solución de CC disponible para el público en general, siendo propiedad de un determinado proveedor que puede comercializar los servicios de cloud que implemente a través de Internet. En este modelo, múltiples clientes acceden a los servicios sobre una infraestructura común, cuya seguridad y disponibilidad es gestionada por el proveedor del servicio. La infraestructura Privada es gestionada integralmente por una organización, independientemente de donde esté alojada, el acceso y configuración están restringidos a sus integrantes. La infraestructura Híbrida está compuesta por algunas de las soluciones previas. Puede ser brindada por diferentes proveedores hacia un conjunto de organizaciones para hacer uso de los servicios disponibles tanto en entornos privados como en públicos. La infraestructura Comunitaria es una infraestructura compartida por varias entidades, brindando servicios a un conjunto específico de usuarios en base a un objetivo común.

Habiendo una amplia disponibilidad de herramientas se pretende determinar una metodología a partir de la cual sea posible la selección de recursos computacionales gestionados con tecnologías de CC en ambientes educativos.

Líneas de Investigación, Desarrollo e Innovación

En el contexto de la investigación para la elaboración del trabajo final de maestría se busca cambiar el paradigma actual en cuanto a la gestión de la información del ámbito académico y administrativo. Para la realización del trabajo mencionado se tendrán en cuenta

como mínimo los trabajos que se mencionan a continuación.

En (P. Mell and T. Grance, 2011) define al CC y sus características en cuanto a los modelos.

En (A. Habbal, S. A. Abdullah, E. O. C. Mkpojiogu, S. Hassan, and N. Benamar, 2017) se describe el diseño y gestión de un modelo de Nube Privada en universidades.

En (S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, and A. Ghalsasi, 2011) se presentan las fortalezas, debilidades, oportunidades y amenazas de la computación en la nube.

En (V. H. Pardeshi, 2014) presenta los desafíos para proporcionar soporte informático para actividades educativas en un entorno de CC.

En (T. Ercan, 2010) se describe el uso efectivo del CC en instituciones educativas.

En (M. Mannir and A. Getso, 2014) se presentan las aplicaciones del CC en instituciones académicas.

En (B. Mohammed and M. Kiran, 2015) se discute los conceptos de virtualización en relación con herramientas de implementación de CC bajo código abierto.

En (H. Chihi, W. Chainbi, and K. Ghdira, 2016) analiza los riesgos de la computación en la nube y lo relaciona con las instituciones de educación superior.

En (A. O. Akande and J.-P. Van Belle, 2016) proporciona una visión general de los conceptos de la computación en la nube para luego centrarse en SaaS.

En (Q. Lei, Y. Jiang, and M. Yang, 2014) realiza una evaluación cuantitativa de plataformas abiertas IaaS utilizando el modelo de referencia presentado por el NIST .

Resultados y Objetivos

Resultados

Los resultados del proyecto son de aplicación directa en la docencia y áreas administrativas de la Facultad de Ciencias Económicas de la Universidad Nacional de Misiones (U.Na.M.)

Se considera que los resultados de la investigación propuesta sean extensivo a las de-

más unidades académicas de la U.Na.M. bajo la modalidad independiente o centralizada dependiendo de los acuerdos con la Unidad Central (Rectorado).

Asimismo, se tiene previsto presentar para su posible publicación los resultados del presente proyecto en revistas y congresos especializados, como así también en conferencias a impartir en diferentes ámbitos.

Objetivo general del proyecto

Desarrollar una metodología válida para identificar conceptos y arquitecturas que permitan evaluar alternativas para la gestión de datos y servicios en la nube, a implementarse en el ambiente académico usando tecnologías de Software Libre. La propuesta permitiría sentar las bases conceptuales para la identificación de recursos computacionales que permita abordar el despliegue de plataformas educativas basadas en el modelo de CC.

Objetivos específicos del proyecto

Efectuar estudios exploratorios de sobre CC, las tecnologías utilizadas, las plataformas de despliegue, los ambientes de utilización, casos éxito, etc., centrando la atención en lo relacionado con el caso de estudio.

Definir las características a considerar para los ambientes académicos que formarán parte de los casos de aplicación de la metodología.

Formación de Recursos Humanos

El equipo de trabajo está integrado por un Doctor y un maestrando con cursadas finalizadas actualmente trabajando en la tesis de maestría con temáticas afines a la del proyecto.

Referencias

A. Habbal, S. A. Abdullah, E. O. C. Mkpogio, S. Hassan, and N. Benamar, "Assessing Experimental Private Cloud Using Web of System Performance Model," *Int. J. Grid High Perform. Comput.*, vol. 9, no. 2, pp. 21–35, Apr. 2017.

P. Mell and T. Grance, "The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology," *Nist Spec. Publ.*, vol. 145, p. 7, 2011.

S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, and A. Ghalsasi, "Cloud computing - The business perspective," *Decis. Support Syst.*, vol. 51, no. 1, pp. 176–189, Apr. 2011.

V. H. Pardeshi, "Cloud Computing for Higher Education Institutes: Architecture, Strategy and Recommendations for Effective Adaptation," *Procedia Econ. Financ.*, vol. 11, pp. 589–599, 2014.

T. Ercan, "Effective use of cloud computing in educational institutions," in *Procedia - Social and Behavioral Sciences*, 2010, vol. 2, no. 2, pp. 938–942.

M. Mannir and A. Getso, "Applications ' of Cloud Computing in Academic Institutions," *Int. J. Inf. Syst. Eng.*, vol. 2, no. 1, pp. 65–72, 2014.

B. Mohammed and M. Kiran, "Analysis of Cloud Test Beds Using OpenSource Solutions," *Proc. - 2015 Int. Conf. Futur. Internet Things Cloud, FiCloud 2015 2015 Int. Conf. Open Big Data, OBD 2015*, pp. 195–203, 2015.

H. Chihi, W. Chainbi, and K. Ghdira, "Cloud computing architecture and migration strategy for universities and higher education," *Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl. AICCSA*, vol. 2016-July, 2016.

A. O. Akande and J.-P. Van Belle, "The Use of Software As a Service by Students in Higher Education Institutions: A Systematic Literature Review," *Int. Conf. Electron. Commer. E-Commerce Smart Connect. World*, pp. 1–6, 2016.

Q. Lei, Y. Jiang, and M. Yang, "Evaluating open IaaS cloud platforms based upon NIST Cloud Computing Reference Model," *Proc. - 17th IEEE Int. Conf. Comput. Sci. Eng. CSE 2014, Jointly with 13th IEEE Int. Conf. Ubiquitous Comput. Commun. IUCC 2014, 13th Int. Symp. Pervasive Syst.*, pp. 1909–1914, 2014.

Nueva propuesta para la administración de recursos y procesos en sistemas distribuidos

Federico Agostini, David L. la Red Martínez, Julio C. Acosta

Facultad de Ciencias Exactas y Naturales y Agrimensura/ Universidad Nacional del Nordeste

9 de Julio 1449, (3400) Corrientes, Argentina,

fagostini@exa.unne.edu.ar, lrm david@exa.unne.edu.ar, julioforever@hotmail.com

Resumen

En los sistemas de procesamiento distribuido es necesario que los procesos que actúan en grupos deban tomar decisiones basados en acuerdos respecto del acceso a recursos; las decisiones pueden estar relacionadas con la realización de determinada actividad que requiera o no la sincronización de los procesos, es decir, que los procesos del grupo estén activos en los mismos lapsos de tiempo en sus respectivos procesadores, requiriendo el uso de recursos compartidos en la modalidad de exclusión mutua mediante consensos estrictos o no.

Así surge el siguiente interrogante: ¿cuáles son los modelos de decisión y los operadores de agregación que habrá que generar para la toma de decisiones en la gestión de grupos de procesos, que trasciendan el enfoque tradicional de las ciencias de la computación, teniendo en cuenta la auto-regulación?

Para ello habrá que considerar diferentes situaciones relacionadas con el hecho de compartir o no recursos y con el nivel de acuerdo o consenso requerido, que podrá ser estricto o no, con posibles requisitos de sincronización.

Los modelos considerarán la posibilidad de imputación de datos faltantes y la fuzzyficación de ciertas variables, utilizando operadores OWA, buscando generar operadores de agregación específicos.

Los modelos desarrollados se evaluarán comparando sus características con los modelos habitualmente utilizados.

Se consideran modelos clásicos para acceder a recursos compartidas en la

modalidad de exclusión mutua utilizando regiones críticas al algoritmo centralizado, al algoritmo distribuido de Lamport, Ricart y Agrawala, al algoritmo de anillo de fichas, entre otros.

Palabras clave: Sistemas operativos - Comunicación entre grupos de procesos - Operadores de agregación.

Contexto

Este estudio se encuadra en el marco del Proyecto de Investigación “Modelos de decisión y operadores de agregación para la administración de procesos en sistemas distribuidos”, acreditado por Resolución N° 241/17 C.S. N° 16F001. 2017-2020, cuyas líneas de trabajo pertenecen al Grupo de Sistemas Operativos y TICs (Res. 725/10 C.D. - FaCENA-UNNE). Además está relacionado con la tesis de maestría “Nueva propuesta para la administración de recursos y procesos en sistemas distribuidos” de la Maestría en Sistemas y Redes de Telecomunicaciones, cuyo plan de trabajo fue aprobado por Res. N° 0733/16.

En los sistemas computacionales de procesamiento distribuido es frecuentemente necesario que los procesos que actúan en grupos deban tomar decisiones basados en el acuerdo; dichos procesos podrán operar en un mismo equipo informático o en varios equipos distribuidos interconectados; las decisiones para las cuales deben alcanzar algún nivel de acuerdo pueden estar relacionadas con la realización de determinada actividad que no requiera el uso de recursos compartidos en la modalidad de

exclusión mutua, o con la realización de determinada actividad que sí requiera el uso de recursos compartidos en la modalidad de exclusión mutua, para lo cual generalmente las exigencias de niveles de acuerdo son mayores que para el caso anterior, pudiendo darse además que los procesos integren grupos que requieran (o no) sincronización (estar activos en sus respectivos procesadores en un mismo lapso de tiempo).

Ante la situación mencionada precedentemente surge el siguiente interrogante: ¿cuáles son los nuevos modelos de decisión que habrá que desarrollar incorporando la perspectiva cognitiva a los modelos clásicos para la toma de decisiones en grupos de procesos, que trasciendan el enfoque tradicional de las ciencias de la computación? Habrá que desarrollar, por lo tanto, los modelos de decisión para la toma de decisiones en grupos de procesos, para los siguientes tipos de situaciones:

- a) Que los procesos accedan a recursos compartidos en la modalidad de exclusión mutua sin constituir grupos de procesos que requieran sincronización (estar activos en sus respectivos procesadores en un mismo lapso de tiempo) y con exigencias estrictas de consenso para lograr el acceso.
- b) Que los procesos accedan a recursos compartidos en la modalidad de exclusión mutua sin constituir grupos de procesos que requieran sincronización (estar activos en sus respectivos procesadores en un mismo lapso de tiempo) y sin exigencias estrictas de consenso para lograr el acceso.
- c) Que los procesos accedan a recursos compartidos en la modalidad de exclusión mutua constituyendo grupos de procesos que requieren sincronización (estar activos en sus respectivos procesadores en un mismo lapso de tiempo) y con exigencias estrictas de consenso para lograr el acceso.
- d) Que los procesos accedan a recursos compartidos en la modalidad de exclusión mutua constituyendo grupos de procesos que requieren sincronización (estar activos en sus respectivos procesadores en un mismo lapso de tiempo) y sin exigencias estrictas de consenso para lograr el acceso.

Introducción

La proliferación de sistemas informáticos, muchos de ellos distribuidos, en los cuales existen múltiples procesos que cooperan para el logro de una determinada función, hace necesario disponer de modelos de decisión que permitan a los procesos intervinientes en los distintos grupos de procesos, tomar decisiones en las que son necesarios diferentes niveles de acuerdo, especialmente cuando se trata del acceso a recursos computacionales compartidos y el sistema debe auto-regular la forma de dicha compartición.

Es especialmente significativo el caso del acceso a las llamadas regiones críticas de memoria por parte de distintos procesos, que pueden estar operando en equipos distribuidos, donde el acceso a las regiones críticas debe hacerse en la modalidad de acceso exclusivo y con el consentimiento de los demás procesos del grupo. Ejemplos de lo mencionado se encuentran en (Tanenbaum, 1996 y 2009), donde se describen los principales algoritmos de sincronización en sistemas distribuidos, en (Agrawal et al., 1991), donde se presenta una solución eficiente y tolerante a fallas para el problema de la exclusión mutua distribuida, en (Ricart et al., 1981), (Cao y Singhal, 2001) y en (Lodha y Kshemkalyani, 2000), donde se presentan unos algoritmos para gestionar la exclusión mutua en redes de computadoras, en (La Red Martínez, 2004), donde se describen los principales algoritmos de sincronización en sistemas distribuidos, en (Stallings, 2005), donde se detallan los principales algoritmos para la gestión distribuida de procesos, los estados globales distribuidos y la exclusión mutua distribuida. Estos temas y otros relacionados también han sido tratados en (Joshi y Holzmann, 2007), (Alagarsamy, 2003), etc.

Los modelos de decisión actualmente disponibles y generalmente aplicables en los sistemas distribuidos se basan en algoritmos de intercambio de permisos que intentan lograr un acuerdo de todos los procesos intervinientes para realizar determinadas

acciones, como el acceso a un área de memoria compartida a la que se debe acceder en la modalidad de exclusión mutua.

Se considera especialmente importante estudiar la aplicación de modelos de decisión para la toma de decisiones en grupo que se desprendan de conceptos cognitivos de la cibernética en general y de la cibernética de segundo orden en particular, en el contexto de sistemas complejos auto-regulados. Se considera en tal sentido que dichos grupos de procesos mejorarían su desempeño mediante los modelos de decisión que se tiene previsto desarrollar incorporando mecanismos de auto-regulación y conceptos de la cibernética de segundo orden en el proceso de toma de decisiones.

Se pretende generar nuevos modelos de toma de decisiones en grupos de procesos distribuidos, contemplando además la aplicación de métodos de imputación de datos para aquellos casos de datos faltantes, por ejemplo, como consecuencia de problemas en las comunicaciones entre los procesos, y fuzzyficación de variables para dar soporte a situaciones donde no es posible o conveniente expresar valores exactos.

Líneas de Investigación, Desarrollo e Innovación

En el contexto del proyecto “Modelos de decisión y operadores de agregación para la administración de procesos en sistemas distribuidos”; iniciado en el 2017 en la UNNE, se busca generar modelos de decisión innovadores y sus correspondientes operadores de agregación para la gestión de grupos de procesos.

En (La Red Martínez, 2004) se describen los principales algoritmos de comunicación en sistemas distribuidos (algoritmos clásicos de las ciencias de la computación).

En (Macedonia et al., 1995) se describe una arquitectura de red para entornos virtuales de gran escala para soportar la comunicación entre grupos de procesos distribuidos.

En (Silberschatz et al., 2006) se presentan los principales algoritmos de coordinación

distribuida y gestión de la exclusión mutua (algoritmos clásicos de las ciencias de la computación).

En (La Red et al., 2011a) se presenta el operador WKC-OWA para agregar información en problemas de decisión democrática.

En (La Red et al., 2011b) se presenta un modelo de decisión en grupo con la utilización de etiquetas lingüísticas y una nueva forma de expresión de las preferencias de los decisores.

En (La Red y Acosta, 2015) se presentan las principales propiedades matemáticas y las medidas de comportamiento relacionadas con los operadores de agregación.

En (La Red y Pinto, 2015) se presenta una revisión acerca de los operadores de agregación, especialmente los de la familia OWA.

En (Chao et al., 2016) se estudia la forma de obtener un vector de prioridades colectivo a partir de diferentes formatos de expresión de las preferencias por parte de los decisores. El modelo puede reducir la complejidad de la toma de decisiones y evitar la pérdida de información cuando se transforman los diferentes formatos en un formato único de expresión de las preferencias.

En (Dong et al., 2016a) se estudia la conexión de la jerarquía lingüística y la escala numérica para el modelo lingüístico de 2-tupla y su uso para tratar información lingüística no balanceada.

En (Dong et al., 2016b) se define un problema complejo y dinámico de toma de decisiones en grupo con múltiples atributos y se propone un método de resolución que utiliza un proceso de consenso para grupos de atributos, de alternativas y de preferencias, presentándose un modelo de decisión para problemas del mundo real.

Resultados y Objetivos

Resultados

Los resultados del proyecto son de aplicación directa en la docencia de la asignatura Sistemas Operativos de la carrera Licenciatura en Sistemas de Información de

la Facultad de Ciencias Exactas y Naturales y Agrimensura de la U.N.N.E.

Se considera que la investigación propuesta es relevante desde el punto de vista social ya que se espera generar el conocimiento científico que permita a posteriori la generación de mejoras en la gestión de recursos compartidos por parte de los sistemas operativos, especialmente de sistemas distribuidos, ampliamente utilizados de manera directa o indirecta por la sociedad en su conjunto.

Asimismo, se tiene previsto presentar para su posible publicación los resultados del presente proyecto en revistas y congresos especializados, como así también en conferencias a impartir en diferentes ámbitos. En tal sentido, ya se han realizado algunas publicaciones, tales como (La Red Martínez, 2017), (La Red Martínez et al., 2017).

Objetivo general del proyecto

Generar modelos de decisión innovadores y sus correspondientes operadores de agregación para la gestión de grupos de procesos.

Objetivos específicos del proyecto

Generar modelos de decisión y sus correspondientes operadores de agregación para la gestión de grupos de procesos que pueden compartir recursos, estudiando la utilización de posibles modificaciones de los operadores de la familia OWA (Yager, 1988, 1993) y la creación de operadores nuevos

Formación de Recursos Humanos

El equipo de trabajo está integrado por un Doctor, una Magister y dos maestrands con cursadas finalizadas. Actualmente uno de ellos se encuentra trabajando en la tesis de maestría con temáticas afines a la del proyecto.

Referencias

Agrawal, D. y El Abbadi, A. (1991). An Efficient and Fault-Tolerant Solution of Distributed Mutual Exclusion. ACM Trans. on Computer Systems. Vol. 9. Pp. 1-20. USA.

Alagarsamy, K. (2003). Some Myths About Famous Mutual Exclusion Algorithms. ACM SIGACT News 34 (3): 94–103.

Cao, G. and Singhal, M. (2001). A Delay-Optimal Quorum-Based Mutual Exclusion Algorithm for Distributed Systems. IEEE Transactions on Parallel And Distributed Systems. Vol. 12, no. 12. Pp. 1256-1268. USA.

Chao, X., Kou, G., Peng, Y. (2016). An optimization model integrating different preference formats, 6th International Conference on Computers Communications and Control (ICCCC), pp. 228 – 231.

Dong, Y., Li, C-C., Herrera, F. (2016a). Connecting the linguistic hierarchy and the numerical scale for the 2-tuple linguistic model and its use to deal with hesitant unbalanced linguistic information, Information Sciences, Volumes 367–368, Pages 259–278, Elsevier.

Dong, Y., Zhang, H., Herrera-Viedma, E. (2016b). Consensus reaching model in the complex and dynamic MAGDM problem, Knowledge-Based Systems, Volume 106, Pages 206–219, Elsevier.

Joshi, R., Holzmann, G. J. (2007), A Mini-Challenge: Build a Verifiable Filesystem, Formal Aspects of Computing, Vol. 19.

La Red Martínez, D.L. (2004). Sistemas Operativos. EUDENE. Argentina.

La Red, D. L.; Doña, J. M.; Peláez, J. I.; Fernández, E. B. (2011a). WKC-OWA, a New Neat-OWA Operator to Aggregate Information in Democratic Decision Problems. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, World Scientific Publishing Company, págs. 759-779. Francia.

La Red, D. L.; Peláez, J. I.; Doña, J. M. (2011b). A Decision Model to the Representative Democracy With Expanded Vote. V. 1 N° 1. Revista Pioneer Journal of Computer Science and Engineering Technology. Pp. 35-45. India.

La Red Martínez, D. L., Acosta, J. C. (2015). Review of Modeling Preferences for Decision Models, Volume 11 – N° 36, European Scientific Journal (ESJ), pp. 1-18, ISSN N° 1857-7881, Macedonia.

La Red Martínez, D. L., Pinto, N. (2015). Brief Review of Aggregation Operators; Volume 22 – N° 4; Wulfenia Journal; pp. 114-137; Austria.

La Red Martinez, D.L. (2017). Aggregation Operator for Assignment of Resources in Distributed Systems; Vol. 8, N° 10; (IJACSA) International Journal of Advanced Computer Science and Applications; pp. 406-419; ISSN N° 2156-5570; The Science and Information (SAI) Organization, England, U.K.

La Red Martinez, D.L.; Acosta, J.C.; Gerzel, S.M.; Rambo, A.R. (2017), New decision making models of processes synchronization in distributed systems; International Conference on Communication and Electronic Information Engineering; ISBN N° 978-94-6252-312-8; pp 65-71; CEIE 2016; Guang-zhou, China; Advances in Engineering Research (AER), Atlantis Press, Vol. 116; 2017.

Lodha, S. and Kshemkalyani, A. (2000). A Fair Distributed Mutual Exclusion Algorithm. IEEE Trans. Parallel and Distributed Systems. Vol. 11. N° 6. Pp. 537-549. USA.

Macedonia, M. R., Zyda, M. J., Pratt, D. R., Brutzman, D. P., Barham, P. T. (1995). Exploiting Reality with Multicast Groups: A Network Architecture for Large-scale Virtual Environments. Proc. of IEEE VRAIS (RTP, NC, Mar., 1995), pp. 2-10.

Ricart, G. y Agrawala, A.K. (1981). An Optimal Algorithm for Mutual Exclusion in Computer Networks. Commun. of the ACM. Vol. 24. Pp. 9-17.

Silberschatz, A., Galvin, P.B. y Gagne, G. (2006). Fundamentos de Sistemas Operativos. 7ma. Edición. McGraw-Hill / Interamericana de España. S.A.U. España.

Stallings, W. (2005). Sistemas Operativos. 5ta. Edición. Pearson Educación S.A. España.

Tanenbaum, A. S. (1996). Sistemas Operativos Distribuidos. Prentice-Hall Hispanoamericana S.A. México.

Tanenbaum, A. S. (2009). Sistemas Operativos Modernos. 3ra. Edición. Pearson Educación S. A. México.

Yager, R. (1988). On Ordered Weighted Averaging Aggregation Operators In Multi-Criteria Decision Making. IEEE Trans. On Systems, Man and Cybernetics 18: 183-190.

Yager, R. (1993). Families Of OWA Operators. Fuzzy Sets and Systems. 59: 125-148.

Visualización Gráfica de resultados de Simulaciones de Redes de Sensores Inalámbricos.

Diego Alberto Godoy^a, Hernán Bareiro^b, Eduardo O. Sosa^c, Lucas.Fiege^d, Juan de Dios Benítez^e, Edgardo A. Belloni^f, Fabián Favret^g

Centro de Investigación en Tecnologías de la Información y Comunicaciones (C.I.T.I.C.)
Departamento de Ingeniería y Ciencias de la Producción-Universidad Gastón Dachary
Av. López y Planes 6519- Posadas, Misiones, Argentina. Teléfono: +54-376-4438677

^a diegodoy@citic.ugd.edu.ar, ^bhbareiro@citic.ugd.edu.ar, ^ceduardo.sosa@citic.ugd.edu.ar,
^dlucas.fiege@citic.edu.ar, ^ejuan.benitez@citic.ugd.edu.ar, ^febelloni@ugd.edu.ar,
^gfabianfavret@citic.ugd.edu.ar

Resumen

En este trabajo se presenta un proyecto de investigación denominado “Simulación en las Tics: Diseño de Simuladores de Procesos de Desarrollo de Software Agiles y Redes De Sensores Inalámbricos para la Industria y la Academia”, cuyo objetivo es diseñar simuladores de procesos de desarrollo de software agiles y de Redes de Sensores Inalámbricos para la Industria y la academia. Particularmente en este artículo se presentan los avances realizados en relación a diseño de una herramienta de visualización gráfica para el análisis de resultados de simulación de WSN.

Palabras claves: Simulación; Visualización Gráfica de Resultados; Redes de Sensores Inalámbricos.

Contexto

El trabajo presentado en este artículo tiene como contexto marco el proyecto de investigación denominado “Simulación en las Tics: Diseño de Simuladores de Procesos de Desarrollo de Software Agiles y Redes De Sensores Inalámbricos para la Industria y la Academia”, registrado actualmente en la Secretaría de Investigación y Desarrollo de la Universidad Gastón Dachary (UGD) con el Código IP

A07003 radicado en el Centro de Investigación en Tecnologías de la Información y Comunicaciones de dicha universidad.

El mismo fue incorporado como proyecto aprobado en el llamado a presentación interna de la UGD de proyectos de investigación N°7 mediante la Resolución Rectoral 07/A/17 y es una continuidad del Proyecto Simulación como herramienta para la mejora de los procesos de software desarrollados con metodologías ágiles utilizando dinámica de sistemas, R.R. UGD N° 18/A/14 y R.R. UGD N° 24/A/15.

Entre las líneas con mayores resultados dentro del proyecto referido, se encuentran las de: Construcción de una plataforma de gestión y simulación de datos de redes de sensores inalámbricos, una interfaz web para el simulador de WSN Shawn, Sistemas de gestión de residuos de la ciudad de Posadas con tecnologías de Internet de la cosas, Sistema de monitoreo de la temperatura en el proceso de secado del Té.

1. Introducción

En la actualidad existen diferentes software de simulación [1] que brindan interfaces visuales con herramientas que permiten establecer parámetros y variables de entorno, como así también

gestionar la configuración de escenarios y realizar combinaciones de todos estos elementos con el objeto de producir diferentes resultados. Pero estas aplicaciones suelen ser para una plataforma específica o bien consumen un alto nivel de recursos para realizar el procesamiento en el equipo cliente. Se puede considerar que NS-2 web [2] [3] es uno de los sistemas similares al propuesto. Respecto a este punto de vista, las aplicaciones Web que emplean una arquitectura cliente-servidor cuentan con una gran ventaja, ya que aligeran la carga computacional de los equipos de los usuarios finales. Principalmente al realizar procesamiento de datos destinados a la generación de resultados gráficos en la simulación.

Otra característica importante de las aplicaciones web es la usabilidad que presentan. Permite simplificar a través de las interfaces gráficas, tareas como la carga de datos relacionados a la configuración previa a la ejecución de la simulación. Inicialmente en el proyecto WEBShawn [4] [5] se realiza un cambio radical en la forma de utilizar el simulador Shawn [2] para WSN, ya que se adaptó el software que originalmente se utilizaba por consola y línea de comandos, para permitir que los usuarios interactúen directamente a través de la Web.

Sin embargo, en el prototipo que se desarrolló no existía una forma de ingresar todos los parámetros de configuración de visualizaciones que ofrece Shawn. Los mismos se debían escribir directamente sobre un archivo de configuración (disponible en la interfaz Web). Al tratarse de múltiples opciones con distintas incidencias en las salidas gráficas, el usuario debía conocer el orden de ingreso y todos los posibles valores válidos que se aceptan por cada parámetro. Lo que dificultaba

al usuario la tarea de ingresar las distintas opciones gráficas.

Otro aspecto que no fue desarrollado en el prototipo inicial de WEBShawn [6] es la gestión de escenarios de simulación, es decir, permitir a los usuarios exportar o importar las configuraciones de los mismos. En este sentido, el simulador Shawn posee también distintos parámetros que se pueden utilizar para dotar a WEBShawn con esta característica.

Es por ello que actualmente se pretende extender las capacidades faltantes de WEBShawn y mejorar las que se puedan realizar con Shawn, aprovechando el proceso para integrar con otras librerías de generación de gráficos. Todo esto tendiente a mejorar específicamente, tanto la interacción como la generación de gráficas de resultados y visualización desde el navegador.

Dada la problemática descrita, se utilizará el prototipo desarrollado en WEBShawn como base para incrementar la funcionalidad del mismo. Para ello se utilizará la adaptación del motor de simulación realizado sobre Shawn, para implementar la biblioteca de visualización VIS [3]. Permitiendo de esta forma desplegar nuevas funcionalidades en cuanto a la visualización de resultados en formato gráfico.

Al continuar con el desarrollo de dicho prototipo, el procesamiento de los datos enviados por el navegador Web y la comunicación con el motor de simulación de Shawn se realizarán en lenguaje PHP. El servidor Web utilizado será Apache sobre GNU/Linux.

El almacenamiento de información relativa a los proyectos de simulación de los usuarios, datos de la aplicación, modelos de simulación y estructuras de

datos, ya se implementaron inicialmente en WEBShawn con el sistema de gestión de bases de datos PostgreSQL. Por lo cual la gestión de nuevos datos de usuario, así también como los datos de sus preferencias de visualización o configuraciones de escenarios serán almacenados utilizando ésta misma tecnología.

La interfaz de usuario que se ejecutará en el navegador Web del lado del cliente, se desarrollará en lenguaje HTML5, JavaScript, CSS y estará disponible para dispositivos tipo PC.

No se realizará verificación de los resultados de simulación con los resultados de una red de sensores inalámbricos en el mundo real, dado que este tipo de revisiones ya fue realizado por los propios autores del software de simulación Shawn.

Se diseñarán tres escenarios de prueba, de modo que pueda demostrarse la versatilidad de WEBShawn en cuanto a visualización de resultados de simulación.

2. Línea de Investigación

Para esta línea se han planteado los siguientes objetivos:

Como objetivo general se propone: Diseñar una herramienta de visualización gráfica de que permita analizar los resultados de simulaciones de WSN en WEBShawn.

Como objetivos específicos se realizarán los siguientes:

- Establecer el estado del arte respecto a simuladores de WSN basados en la Web.

- Analizar las alternativas de visualización soportadas por Shawn y la gestión de escenarios de simulación en el mismo.
- Especificar los requerimientos de datos para captar los parámetros de visualización deseados y la gestión de escenarios de simulación.
- Construir un prototipo de aplicación Web que incluya los módulos de software necesarios para gestionar escenarios de simulación en WEBShawn, tanto el componente front-end como el back-end de la aplicación.
- Definir tres escenarios de prueba donde se comprobarán las posibilidades de visualización.
-

3. Resultados

Como parte del proceso de adaptación de los nuevos parámetros que se ingresan desde el cliente, se implementó un cambio generalizado en toda la interfaz gráfica. Para ello se utilizó la biblioteca Bootstrap [4]. En la Figura 1 se puede observar parte de la nueva interfaz Web realizada en Bootstrap, específicamente un formulario que permite ingresar los parámetros de configuración gráfica de los nodos de la red. Una vez que el usuario configura y corre una simulación con determinados parámetros, se agregó también la posibilidad de guardar y cargar dicha configuración las veces que se necesite a fin de poder comparar con otras corridas que se realicen, utilizando otros parámetros.

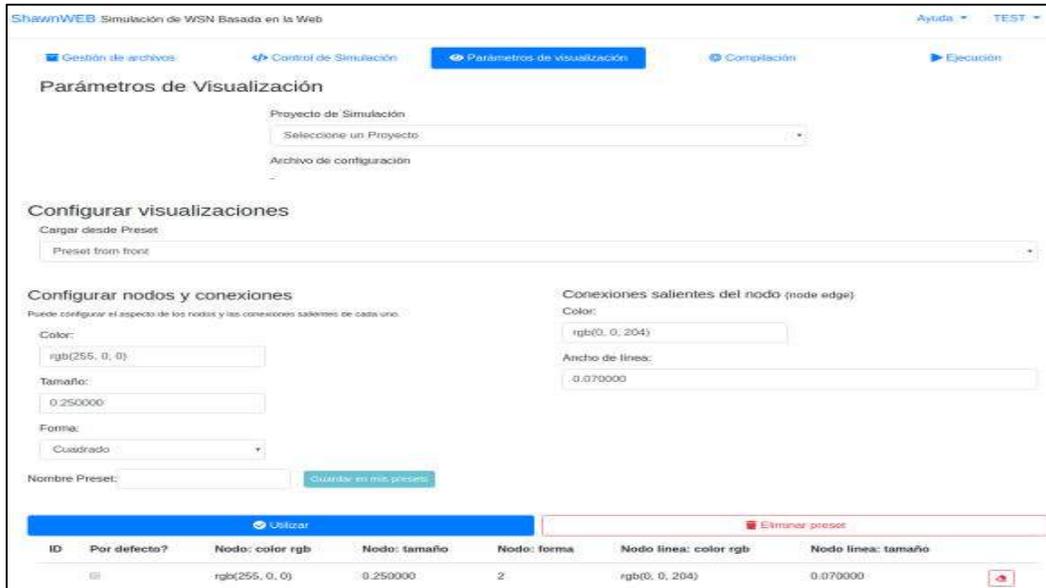


Figura 1. Configuración del aspecto gráfico de los nodos

Una vez que el usuario configura y corre una simulación con determinados parámetros, se agregó también la posibilidad de guardar y cargar dicha configuración las veces que se necesite a fin de poder comparar con otras corridas que se realicen, utilizando otros parámetros.

Continuando con la interfaz Web, en lo que se refiere a la visualización de resultados, se desarrolló un módulo completo que permite generar gráficos con la ubicación de los nodos y la interconexión de los mismo. Utilizando la biblioteca VIS se puede generar la visualización gráfica de los parámetros previamente ingresados como ser color, tamaño, forma y etiquetas con datos de los nodos. En la Figura 2 se puede observar una salida gráfica generada con VIS.

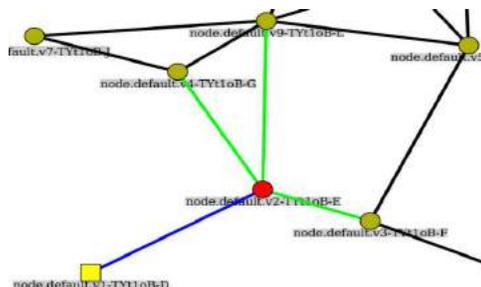


Figura 2. Salida gráfica generada con la biblioteca VIS.

En este ejemplo generado luego de correr una simulación, se determinó que el nodo gateway tenga forma cuadrada y color amarillo. Los nodos que estén conectados directamente (a un solo salto) al nodo gateway tengan forma circular, un color rojo y la interconexión quede trazada en color azul. Para el resto de los nodos se determinó la misma forma pero con color verde e interconexión con líneas en color negro. Además todos los nodos que tengan conexión de hasta dos saltos al gateway se trace dicha conexión en color verde claro.

Como se puede apreciar este tipo de salidas gráficas permite visualmente y de forma ágil analizar diferentes aspectos de una determinada red de nodos sensores.

Del lado del servidor, las principales adaptaciones se hicieron con el objetivo de que todo el prototipo sea más portable a nivel Sistema Operativo. Para ello se creó una imagen completa de WEBShawn utilizando el software contenedor de aplicaciones Docker [5]. De esta forma cualquier usuario que quiera instalar el servidor completo de WEBShawn solo tiene que instalar previamente Docker y con un comando

replicar el servidor configurado con todas las herramientas de software instaladas listas para utilizar. Esto incluye el servidor Web Apache, PHP, la base de datos PostgreSQL, el simulador Shawn y todas sus dependencias.

4. Formación de Recursos Humanos

El equipo de trabajo se encuentra formado por cuatro investigadores con distintos niveles de posgrado, un Doctor en Ciencias Informáticas y Magister en Redes de Datos; un Doctor en Tecnologías de la Información y Comunicaciones Magister y Especialista en Ingeniería de Software; un Maestrando de Ingeniería de la Web; dos Maestrando en Redes de Datos y ocho estudiantes en período de realización de trabajos finales de grado en el contexto de las carreras de Licenciatura en Sistemas de Información y de Ingeniería en Informática de la UGD. Actualmente, el número de tesinas de grado aprobadas en el contexto de este proyecto, es de cinco, y otras tres en proceso de desarrollo. El número de tesis de maestría terminadas relacionadas con este proyecto es de una.

5. Bibliografía

- [1] Jerry Banks, *Discrete-event system simulation*. New Jersey: Pearson, 2010.
- [2] Barun Kumar Saha. NS2Web. [Online]. <http://vlssit.iitkgp.ernet.in/>
- [3] Information Sciences Institute. The Network Simulator 2. [Online]. <https://www.isi.edu/nsnam/ns/>
- [4] Diego Alberto Godoy, Eduardo Omar Sosa, Hernán Bareiro, and Rebeca Díaz Redondo, "Redes de Sensores Inalámbricos: Interfaz Web para Shawn," in *XX Congreso Argentino de Ciencias de la Computación*, Buenos Aires, 2014.
- [5] Diego Alberto Godoy, Eduardo Omar Sosa, Rebeca Pilar Díaz Redondo, and Santiago Hernán Bareiro, "WebShawn, simulating wireless sensors networks from the web," in *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Roma, 2017, pp. 190-195.
- [6] Stefan Fischer, Dennis Pfisterer, and Sándor P. Fekete, "Shawn: The fast, highly customizable sensor network simulator," *Braunschweig University of Technology and University of Lubeck*, no. ISBN 1-4244-1231-5, Alemania 2007.
- [7] Dennis Pfisterer and Marvin Frick. Shawn Wiki Visualization GitHub. [Online]. <https://github.com/itm/shawn/wiki/Visualization>
- [8] Bootstrap. [Online]. <https://getbootstrap.com/>
- [9] Docker. [Online]. <https://www.docker.com/>

IMPLEMENTACIÓN DE SISTEMAS DE CONTROL AUTOMÁTICO PARA CULTIVOS HIDROPÓNICOS EN INVERNADEROS DE LA PROVINCIA DE MISIONES

Marcelo Marinelli, Graciela Lombardo, Kuna Horacio, Guillermo Wurn, Ruben Urquijo, Veronica Gonzalez

Departamento de Informática, Facultad de Ciencias Exactas Químicas y Naturales.
Universidad Nacional de Misiones
Felix de azara 1552, Posadas, Misiones
Te:376-4422186

marcelomarinelli@fceqyn.unam.edu.ar, gracielalombardo@gmail.com, hdkuna@gmail.com
guillermowurn@gmail.com, chinourquijo@gmail.com, gonzalezveronica187@gmail.com

RESUMEN

El objetivo de esta línea de investigación es desarrollar, implementar y evaluar sistemas de control de distintos dispositivos de cultivos hidropónicos en invernaderos de la provincia de Misiones, utilizando sistemas embebidos y minicomputadores de placas de hardware libre o de bajo costo. El software empleado en el sistema utilizará técnicas de inteligencia artificial para evaluar las variables del proceso y generar señales de control. Se utilizarán distintos tipos de sensores de humedad, temperatura, conductividad eléctrica y pH. También se implementarán sistemas de telemetría basado en “Internet de las Cosas” (IoT, Internet of Things) de manera de poder advertir en forma remota, a través de mails, SMS o redes sociales, anomalías en el estado de las variables de los sistemas que intervienen en los cultivos.

CONTEXTO

Este proyecto se enmarca en el “Programa de Investigación en Computación” del Centro de Investigación y Desarrollo Tecnológico de la Facultad de Ciencias Exactas Químicas y Naturales de la Universidad Nacional de Misiones; también vinculado con el Doctorado en Ciencias Aplicadas y la Maestría en Tecnologías de

la Información de la Universidad Nacional de Misiones.

Dentro del proyecto se desempeñan docentes, tesisistas y becarios de las carreras de Analista en Sistemas de Computación, Licenciatura en sistemas de Información, Profesorado en Física, Maestría en Tecnologías de la Información y Doctorandos en Ciencias Aplicadas de la Facultad de Ciencias Exactas, Químicas y Naturales de la Universidad Nacional de Misiones

1. INTRODUCCIÓN

Los productores de hortalizas de la Provincia de Misiones utilizan métodos de siembra directa para la obtención de plantines y su posterior trasplante en canteros. Este sistema produce un importante desperdicio del rendimiento de germinación de las semillas y una reducción del tamaño de los plantines que compiten entre sí dada su proximidad.

El cultivo sin tierra, denominado hidropónico, es una alternativa a la producción tradicional que tiene ventajas en lo que hace al proceso de producción de hortalizas, debido a que se puede tener un mayor control de las variables que afectan al proceso de desarrollo y crecimiento de las plantas, evita la posible contaminación producida por el suelo y además se tiene

preciso control de los nutrientes necesarios para una producción exitosa.

El proyecto propone desarrollar e implementar sistemas hidropónicos automatizados en invernaderos, por medio de evaluación y control de las variables intervinientes en el proceso. Se utilizarán técnicas de control de procesos basadas en inteligencia artificial y sistemas embebidos de bajo costo.

Del relevamiento bibliográfico surgió que, habiendo hecho una comparación entre los distintos sistemas de hardware de microcomputadoras de placa reducida (Single Board Computer o SBC), se optó por la utilización de Raspberry Pi 2 [1]. Esta SBC tiene la característica de ser un dispositivo de bajo costo y soporta distintos sistemas operativos como GNU/Linux ARM (Debian, Fedora, Arch Linux), RISC OS2.

1.1 Diseño de sistemas de control con hardware



Figura 1. Raspberry Pi 2 modelo B.

La placa utilizada, Raspberry Pi 2 modelo B (Figura 2), tiene un tamaño reducido (85 mm de longitud por 56 mm de ancho) y posee una salida HDMI, una ethernet y 4 USB con lo que si se conecta un monitor, teclado y ratón se convierte en un ordenador. Además, posee una CPU ARM1176JZF-S (armv6k) a 700 MHz3, GPU Broadcom VideoCore IV3, memoria 512 MiB, capacidad de almacenamiento tarjeta SD o SDHC.

Puertos GPIO

Posee un total de 17 puertos configurables como entrada o salida, y por defecto están configurados como entradas, excepto los GPIO 14 y 15, que operan como salidas (Figura 2).

En este caso se utilizó el puerto GPIO 5 como entrada digital, para obtener los datos provenientes del sensor de humedad y temperatura.

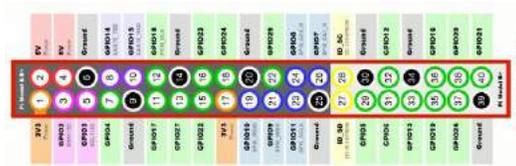


Figura 2. Salidas GPIO de Raspberry Pi.

1.2. Sensor de temperatura DHT22

Para obtener los valores de las variables humedad relativa y temperatura, se utilizaron sensores DHT22 (Figura 3 y Tabla 1). Estos sensores poseen un rango de operación de 0 a 100 % HR y de -40 a 80 °C y cuentan con una precisión de humedad de 2% RH y de temperatura 0,5 %. Por otra parte, tienen la característica de trabajar con protocolo serial enviando los datos por una sola vía, por el pin de datos (Figura 4).

Tabla 1. Pines de salida de DHT22.

Pin	Función
1	5 v
2	Datos
3	NC
4	GND



Figura 3. Sensor DHT22.

De acuerdo con la Tabla 1, el pin 1 corresponde a la alimentación, el pin 4 a masa (GND) y el pin 2 tiene una resistencia de pull up de 10 K, por donde son enviados los datos de humedad y temperatura en forma serial hacia el puerto GPIO, correspondiente de la Raspberri Pi (Figura 4).

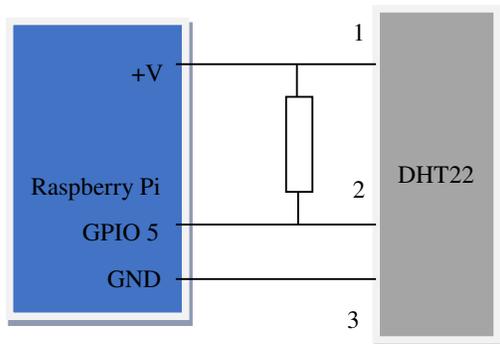


Figura 4. Conexiones del sensor DHT22 con el puerto GPIO.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se está trabajando en distintas líneas de investigación y desarrollo, relacionadas con la automatización de sistemas de cultivos hidropónicos e invernaderos, utilizando sistemas embebidos y técnicas de IA y de IoT., a continuación, se detallan las más importantes:

- Sistemas de control de cultivos hidropónicos con sistemas embebidos de bajo costo.

- Desarrollo de controladores difusos para el control ambiental de invernaderos.
- Telemetría con tecnología bluetooth para el control de procesos y variables ambientales.
- Desarrollo de aplicaciones IoT para telemetría de variables ambientales, provenientes de invernaderos mediante el uso de email y twitter.

3. RESULTADOS OBTENIDOS Y ESPERADOS

En el marco de este programa se concluyó con la tesis doctoral **“Control Automatizado de cultivos Hidropónicos mediante Lógica Difusa”** (Marinelli, 2015), del Doctorado en Ciencias Aplicadas de la Universidad Nacional de Misiones. También se concluyeron dos tesis de grado de licenciatura en Sistemas de Información: **Desarrollo de un Sistema de Navegación y Telemetría en Tiempo Real para Un Robot Movil desde un Smart Phone Via Wifi (IEEE 802.11)**, Autor: Cichanowski, Miguel Alejandro y **Desarrollo e implementación de un controlador difuso para la navegación de robots móviles utilizando el algoritmo de Wang & Mendel**, Autor: Lisandro Solonezen.

Se realizaron las siguientes publicaciones con referato:

- Marinelli, M., Acosta, N., Toloza, J. M., & Kornuta, C. (2016). Control difuso de una cámara de germinación. In XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016).
- Marcelo Marinelli, Guillermo Wurm. (2016). Sistema de adquisición en tiempo real de conductividad eléctrica y pH en nutrientes hidropónicos. IV

- Seminario Argentina-Brasil de Tecnologías de la Información y la Comunicación (SABTIC 2016). ISBN 978-987-3619-15-1.
- XXII CACIC2016 (Congreso Argentino De Ciencias De La Computación) 3 al 7 de octubre de 2016 U.N.S.L, San Luis.
 - IV Seminario Argentina-Brasil de Tecnologías de la Información y la Comunicación (SABTIC 2016). 4 y 5 de noviembre de 2016, UNNE, Corrientes.
 - Marinelli, M., & Urquijo, R. (2017). Sistema de control de una cámara de germinación hidropónica con IoT. In XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017).
 - Marinelli, M., Acosta, N., Toloza, J. M., & Kornuta, C. (2017). Fuzzy Control of a Germination Chamber. *Journal of Computer Science & Technology*, 17.
 - Lombardo, G. C., & Marinelli, M. (2017). Uso de controladores difusos en el proceso de evaluación en matemática. In XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017).
 - Marinelli, M., Lombardo, G., Kornuta, C., Wurn, G., Solonezen, L., & Cichanowski, M. (2017, August). Automatización de sistemas de cultivos hidropónicos. In XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires).

Como resultados esperados se desarrollará un sistema de control ambiental de invernaderos con telemetría basada en IoT. También sistemas el control de las variables ambientales y actuadores en el proceso de control de cultivos hidropónicos. Permitiendo integrar la representación de los datos, los sistemas de telemetría y comunicación de alarmas en forma integrada con los servicios de Internet.

4. FORMACIÓN DE RECURSOS HUMANOS

En este proyecto se están desarrollando una tesis licenciatura en Sistemas de Información y una tesis de la Maestría en Tecnologías de la Información, ambas carreras de la Facultad de Ciencias Exactas, Químicas y Naturales de la Universidad Nacional de Misiones.

Se prevén incorporar tesis de la Maestría en Tecnología de Información y del Doctorado en ciencias Aplicadas de la de la misma casa de altos estudios citada en el párrafo precedente.

5. BIBLIOGRAFÍA

1. Pi—Teach, R. (2016). learn, and make with Raspberry Pi. *Raspberry Pi [Internet].*[cited 23 Mar 2017]. <https://www.raspberrypi.org>.
2. Jain, S., Vaibhav, A., & Goyal, L. Raspberry Pi based interactive home automation system through E-mail. In *Optimization, Reliability, and Information Technology (ICROIT), 2014 International Conference on* (pp. 277-280). IEEE. (2014)
3. Kleinfeld, R., Steglich, S., Radziwonowicz, L., & Doukas, C. glue. things: a Mashup Platform for wiring the Internet of Things with the Internet of Services. In *Proceedings of the 5th International Workshop on Web of Things* (pp. 16-21). ACM. (2014)

4. Marinelli, Marcelo J, Otegui, Mónica B, Zapata, Pedro D, & Acosta, Nelson H. Control Difuso de una Cámara de Germinación para Hidroponía. *Revista de Ciencia y Tecnología*, (24), 42-47. Recuperado en 25 de abril de 2017, de http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1851-75872015000200007&lng=es&tlng=es.
5. Takagi, T., & Sugeno, M. Fuzzy identification of systems and its applications to modeling and control. *IEEE transactions on systems, man, and cybernetics*, (1), 116-132. (1985)

El problema de las comunicaciones rurales: estudio y selección de las mejores soluciones

Antonio Castro Lechtaler^{1,2,3 y 4}; Alejandro Arroyo Arzubi¹; Fernanda Carmona³;
Antonio Foti⁴; Rubén Fusario¹; Anibal Intini¹; Alejandro Oliveros⁴;
Alejandro Echazú¹; Germán Kurt Grin².

¹ Universidad de la Defensa, Facultad del Ejército, Escuela Superior Técnica, Laboratorio de Redes (RedLab); Buenos Aires, C1426; ² Universidad de Buenos Aires, Facultad de Ciencias Económicas, Instituto de Investigaciones en Administración, Contabilidad y Métodos Cuantitativos para la Gestión – (IADCOM/CISTIC), Buenos Aires, C1120; ³ Universidad Nacional de Chilecito, Chilecito, Provincia de La Rioja, F5360; ⁴ Universidad Nacional de 3 de Febrero, Sáenz Peña, Provincia de Buenos Aires, B1674.

antonio.castrolechtaler@gmail.com; aarroyo_arzubi@hotmail.com; fbcarmona69@gmail.com;
foti.antonio@gmail.com; aintini@est.iue.edu.ar; rfusario@gmail.com; aoliveros@untref.edu.ar;
alejandroechazu@yahoo.com; german.grin@gmail.com;

RESUMEN

En muchas regiones del mundo la llegada de servicios de comunicaciones a comunidades pequeñas, con baja densidad poblacional, ha sido un problema que en muy pocos casos ha sido resuelto, y en muchos de ellos a costos elevados.

Las razones son obvias. Éstas carecen de interés comercial para que las empresas que brindan los servicios públicos de comunicaciones quieran brindar estas prestaciones. El reciente crecimiento de la teledensidad¹ en las zonas urbanas, impulsado por la tecnología móvil, ha hecho que la brecha digital entre las zonas rurales y urbanas se haya ampliado [1].

En muchas regiones y países, las actividades rurales tienen una importancia significativa en la economía. Además, la falta de estos servicios impide a estos grupos poblacionales acceder a una educación acorde con sus necesidades, impide en muchos casos contar con una adecuada atención de la salud en casos de urgencia y resiente la actividad económica al impedirle conocer el valor de sus productos en tiempo y forma.

Es por ello, que en muchos países se va produciendo una despoblación de las zonas

rurales y un desplazamiento de ellas a las grandes urbes provocando todo tipo de problemas sociales.

En busca de soluciones económicamente factibles, utilizando nuevas tecnologías ya que existen en el mercado se ha formado un *Grupo de Investigación organizado en Red de Universidades Nacionales*, que tiene por objetivo buscar una solución a este problema de la falta de conectividad en dichas zonas, con el objeto de que se puedan brindar servicios isócronos y de datos de banda ancha con acceso a la Red Internet.

La idea central de la investigación es buscar distintas alternativas, que seguramente diferirán de las utilizadas en los países centrales, pero que pueden constituir una solución a este problema.

En el caso particular de Argentina las distancias son condicionantes por su importancia, y la densidad poblacional es sustancialmente muy inferior a la que se puede encontrar en otros países, especialmente los desarrollados.

Es por ello que se orientó el estudio hacia las técnicas digitales inalámbricas *wireless technologies* -en especial aquellas de largo alcance tales como: microondas, 802.11, WiMax, CDMA450, 802.22 y otras similares- podrían dar solución al problema planteado.

Palabras Clave:

¹ Se entiende por teledensidad a “la cantidad de teléfonos fijos más los móviles en uso por cada 100 personas que viven dentro de un área”. Una teledensidad superior a 100 significa que hay más teléfonos que personas. Los países en vías de desarrollo pueden tener una teledensidad de menos de 10.

CSMA/CA, WLAN, 802.11, 802.22, TVWS.

CONTEXTO

Este Grupo de Investigación trabaja, los problemas de conectividad que se presentan en las comunidades rurales, con el apoyo de Entes Nacionales que están vinculados con esta problemática, en particular, en comunidades muy pequeñas que carecen de comunicaciones de banda ancha u otras de tamaño algo mayor que cuentan con facilidades, pero que hasta el momento son claramente insuficientes para el desarrollo de actividades rentables o para participar adecuadamente en Redes Sociales.

Estas áreas son claramente no rentables para las empresas proveedoras de servicios de telecomunicaciones y consecuentemente no son de interés comercial para esas empresas, las cuales no invierten en el desarrollo de la infraestructura adecuada.

Los trabajos que se realizan con el grupo de investigadores tienen como objetivo principal hallar soluciones técnicas, a costos razonables, para los problemas de conectividad descriptos.

El Grupo tiene su sede en el Laboratorio de Redes – **RedLab**, de la Escuela Superior Técnica “Gral. Div. Manuel N. Savio” (EST) de la Universidad de la Defensa (UNDEF) y simultáneamente, en las Universidades de Buenos Aires (UBA) y las Nacionales de Chilecito (UNdeC) y Tres de Febrero (UNTREF).

Los trabajos realizados hasta el presente han incluido pruebas de campo, en base a subsidios obtenidos por parte de la Agencia Nacional de Promoción Científica y Tecnológica [2] y fondos aportados por las universidades participantes.

Recientemente, y mediante un concurso público con jurados externos, la Universidad de la Defensa Nacional otorgó a este proyecto un Subsidio Especial [3] de 100.000,00 \$ para continuar con su ejecución durante el año 2018.

En este proyecto, se continúan realizando variadas pruebas efectuadas, utilizando diversas tecnologías existentes en el mercado.

En particular, se está trabajando sobre la base de analizar y probar las posibilidades que brindan los equipos que responden a la Recomendación 802.22 de la IEEE; y desde el punto de vista de su posible implementación en las instalaciones existentes que posee el Sistema Argentino de Televisión Digital de la Empresa ARSAT. Actualmente se habla también de las tecnologías denominadas **TV White Spaces** – **TVWS** que utilizan las frecuencias de televisión que se encuentran libres, para llevar Internet de banda ancha a zonas apartadas, funcionando armónicamente con los canales de televisión adyacentes sin generar ningún tipo de interferencias.

Por otra parte, cabe manifestar, que han manifestado por escrito su interés en estas investigaciones en carácter de Entidades Adoptantes, el Instituto de Investigaciones Científicas y Técnicas para la Defensa - CITEDEF, y el Consejo Profesional de Ingeniería de Telecomunicaciones, Electrónica y Computación - COPITEC, que de esta manera han dado su aval para este proyecto, dado su interés en ser aplicado en beneficio de los productores rurales y de sus propios sistemas y redes.

1. INTRODUCCIÓN.

El problema de las comunicaciones rurales ha movilizado a distintos grupos de investigación y empresas de telecomunicaciones a buscar soluciones a este tipo de problemas, por cuanto estas zonas geográficas proporcionan cantidades significativas de productos alimenticios en sus diferentes etapas de fabricación y constituyen una trascendente fuente de productos básicos de exportación e ingresos de divisas.

En muchos países participan generando un porcentaje significativo del producto bruto interno de ellos.

La serie de Recomendaciones 802.XX incluye un conjunto de normas que regulan el funcionamiento de las comunicaciones inalámbricas.

Luego de evaluar con resultados poco satisfactorios los equipos que utilizaban la norma 802.11 fueron apareciendo distintas

tecnologías que permitieron ampliar esta Serie merced al trabajo de distintos grupos de investigación [4-8].

Estos trabajos culminaron el 1 de julio de 2011 cuando finalizó el proceso de aprobación, del estándar “IEEE 802.22 - “IEEE 802.22: Cognitive Wireless Regional Area Network - Medium Access Control (MAC) and Physical Layer (PHY). Specifications: Policies and Procedures for Operation in the TV Band²”.

La misma fue aprobada con el apoyo del Comité LAN/MAN³ de la IEEE [9].

Este nuevo estándar proporciona una opción que permite establecer enlaces inalámbricos full dúplex a distancias de entre 30 a 70 km entre antenas, utilizando frecuencias no restringidas por las regulaciones gubernamentales.

La norma que pertenece a la serie 802.XX⁴ tiene por objeto establecer los criterios para el despliegue de múltiples productos interoperables de la misma, ofreciendo acceso a la banda ancha fija en diversas áreas geográficas, incluyendo especialmente los de baja densidad de población en las zonas rurales, y evitar la interferencia a los servicios que trabajan en la televisión de radiodifusión. La misma es conocida actualmente como Red Inalámbrica de Área Regional y está pensada para operar principalmente como una forma de poder acceder a servicios de banda ancha a redes privadas de datos ubicadas en Zonas Rurales.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO.

Las líneas de Investigación y desarrollo son las siguientes:

2.1. Estudio de todas las Recomendación de la Serie 802.X para determinar la viabilidad de su uso en las comunicaciones rurales.

2.2. Estudio de casos concretos de uso actual de equipamientos que utilicen la Norma

802.22 para comunicaciones en distancias de hasta 100 km.

2.3. Ventajas y posibilidades del uso de los Espacios Blancos que existen en el ancho de banda que se utiliza para los Sistemas de Televisión Digital Terrestre.

2.4. Posibilidad de aprovechamiento de las instalaciones de las estaciones de Televisión Digital Terrestre instaladas a lo largo del país, para la instalación del equipamiento necesario para el funcionamiento de los equipos necesarios para brindar comunicaciones rurales.

2.5. Estudio práctico y teórico [10, 11], de las interferencias entre canales debido al uso intensivo de las comunicaciones inalámbricas para todo tipo de servicios de comunicaciones.

2.6. Determinación de la posible utilización de las frecuencias asignado para la transmisión del Sistema Nacional de Televisión Abierta para combinarlo con un sistema basado en estas normas, para la utilización en las comunicaciones rurales utilizando los espacios blancos.

2.7. Estudio de la posibilidad de utilizar TVWS como tecnología emergente. Ésta ya tiene numerosos productos desarrollados en el mercado con precios muy competitivos.

3. RESULTADOS OBTENIDOS / ESPERADOS.

3.1. Se han obtenido los siguientes resultados:

3.1.1. Del estudio del estándar IEEE 802.22 y del equipamiento analizado se ha determinado que tanto por el alcance como por las prestaciones puede resultar una solución a las comunicaciones rurales, en las condiciones descriptas.

3.1.2. El mismo está teóricamente dentro de las distancias requeridas para satisfacer los objetivos del proyecto.

3.1.3. Las frecuencias asignadas al espectro de la Televisión Digital Abierta en la modalidad de radio cognitiva [12, 13] son adecuadas a las características de este proyecto.

3.1.4. Se ha considerado que el Sistema de TDA puesto en marcha puede ser una oportunidad para que esta norma sea

² “IEEE 802.22 - Red de Área Regional Cognitiva Inalámbrica de Control de Acceso al Medio MAC) y la Capa Física (PHY). Especificaciones, Políticas y Procedimientos para la Operación en las Bandas de Televisión”.

³ LAN: Local Area Network; MAN: Metropolitan Area Network.

⁴ Redes Inalámbricas.

considerada en la reasignación del espectro, actualmente en estudio, por parte de la Comisión Nacional de Comunicaciones.

3.1.5. Al trabajar en las mismas frecuencias que otros servicios, pero protegiendo las transmisiones de los operadores principales habilitados en las mismas, el estándar posee un conjunto de capacidades que incluyen:

- Detección del espectro;
- Servicios de geolocalización;
- Acceso a base de datos con información sobre el estado del espectro;
- Registro y seguimiento de la gestión del conjunto de los canales que están operando en un determinado momento en una zona geográfica determinada [14].

Estas capacidades lo hacen también que pueda ser una solución al problema planteado.

3.2. En lo que respecta a los resultados esperados se estima ellos podría sintetizarse en los siguientes:

3.2.1. La 802.22 podría estar capacitada para:

- Explotar y detectar canales operativos que podrían producir interferencias tales como: transmisiones de televisión;
- la emisión de micrófonos inalámbricos;
- Las transmisiones de dispositivos de protección como podrían ser faros inalámbricos u otras transmisiones como por ejemplo la telemetría médica (que requiere ser protegida por la autoridad regulatoria local).

3.2.2. La capa de enlace toma elementos de la norma 802.3, de amplia difusión y probada eficiencia. Estas características deben ser verificadas mediante trabajos de campo.

3.2.3. Se estima que la existencia de un sistema de televisión por radiodifusión ya instalado en un gran porcentaje evitaría tener que usar una porción adicional del espectro de frecuencias, cada vez más escaso y congestionado.

3.2.4. Se analizará y se buscarán resultados sobre la utilización de los espacios Blancos con el objeto de buscar reducir el uso del Espectro de Frecuencias.

Como resultado de estas actividades se presentó un trabajo [15] en el XXII Congreso Argentino de Computación - CACIC 2016, el que *fue seleccionado para ser publicado en*

el libro de los mejores artículos que se publica anualmente, en este caso del CACIC 2016 (31 trabajos) [16], en base a los resultados realizados por los evaluadores sobre el total de artículos presentados.

Se estima que se deberá continuar con el estudio de esta recomendación en sus aspectos técnicos, para determinar fundamentalmente sus limitaciones, si ellas existieran, todo ello sin perjuicio de buscar otras opciones.

3.2.5. Se efectuará un relevamiento del equipamiento que el mercado está ofreciendo sobre esta norma y un análisis de las capacidades del mismo; buscando obtener una idea de las capacidades y costos de este tipo de equipamientos para cubrir distintas zonas del territorio nacional.

3.2.6. Se continuarán las actividades de campo para verificar el verdadero rendimiento del equipamiento y la dificultad que requerirá su despliegue, tal como el equipo de investigación efectuó sobre el terreno el Proyecto Corral de Lorca.

Las instalaciones de las estaciones base del sistema de televisión digital terrestre, instaladas sobre *shelters*, sin duda pueden ser útiles para el despliegue de parte de los equipos requeridos por la 802.22.

4. FORMACIÓN DE RECURSOS HUMANOS.

Desde el año 2015 en este grupo trabajan Investigadores en Formación y alumnos de las carreras de grado y posgrado vinculadas con los temas que hacen tanto a las comunicaciones, como a la seguridad de los sistemas que podrían ser utilizados.

Durante el año 2017 y en este año 2018, se han sumado al proyecto nuevos investigadores en formación y alumnos de las distintas universidades participantes en especial de las Carreras de Ingeniería en Informática y Electrónica.

Algunos de ellos han recibido las becas *Estímulo a las Vocaciones Científicas*, perteneciente al *Programa Estratégico de Investigación y Desarrollo, Plan de Fortalecimiento (Componente de Formación*

de Recursos Humanos) del Consejo Interuniversitario Nacional.

Cabría la posibilidad incluso algunos de ellos, realicen su Trabajo Final de Carrera en algún tema de los que aborda la presente línea de investigación.

Los integrantes son docentes y alumnos de las asignaturas las siguientes asignaturas en las distintas Universidades participantes: Tecnología de las Comunicaciones; Sistemas de Comunicaciones I y II; Comunicaciones Inalámbricas; Redes de Computadoras, entre otras.

5. BIBLIOGRAFÍA.

- [1] <https://www.itu.int/en/ITU-D/Technology/Pages/RuralCommunications.aspx>
- [2] Proyecto FONCyT - ANPCyT. PICTO 11-18621. Redes Privadas Comunitarias. Proyecto finalizado y aprobado. Antonio Castro Lechtaler (Director).
- [3] Resolución Rectoral UNDEF N° 282/2017 de fecha 27 de octubre de 2017. Expediente N° 606/2016.
- [4] García Guibout, J., García Garino C., Castro Lechtaler, A. R., Fusario R., y Sevilla, G., (2007) Physical and Link Layer in Power Line Communications Technologies. *Proceedings of 13th of Argentine Congress on Computer Science*. ISBN 978 - 950 - 656 - 109 - 3. Pág. 56 a 67.
- [5] García Guibout, G., García Garino, C., Castro Lechtaler, A. R., Fusario R. y Sevilla, G. (2007) Power Line Communications in the Electric Network. *Proceedings of 13th of Argentine Congress on Computer Science*. ISBN 978 - 950 - 656 - 109 - 3. Pág. 68 a 79.
- [6] García Guibout, J., García Garino. C., Castro Lechtaler, A. y Fusario, R., (2008). Transmission voice over 802.11. *Proceedings of 14th of Argentine Congress on Computer Science*. ISBN 978 - 987 - 24611 - 0 - 2. Pág. 307 a 318.
- [7] Castro Lechtaler, A., Foti, A., Fusario, R., García Garino, C., y García Guibout, J., (2009) Communication Access to Small and Remote Communities: The Corral de Lorca Project. *Proceedings of 15th of Argentine Congress on Computer Science*. ISBN 978 - 897 - 24068 - 4 - 1. Pag. 1.117 a 1.126.
- [8] A. Castro Lechtaler, A. Foti, C. García Garino, J. García Guibout, R. Fusario and A. Arroyo Arzubi. Proyecto Corral de Lorca: Una solución de conectividad a grupos poblacionales pequeños, aislados y distantes de centros urbanos. *Proceedings de la Novena Conferencia Iberoamericana en Sistemas, Cibernética e Informática: CISCI 2010 - Volume III - ISBN - 13: 978 - 1 - 934272 - 96 - 1. pp. 121 a 127. Orlando, USA. June 2010.*
- [9] IEEE 802.22 - Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) *Specifications Policies and Procedures for Operation in the TV Bands*.
- [10] Gómez, C., (2013). Spectrum Regulation and Policy Officer Radiocommunication ITU. Apia, Samoa. www.itu.int/ITU-D/asp/CMS/Events/.../ITU-APT-S3_Cristian_Gomez.pdf
- [11] CEPT Report 24. A preliminary assessment of the feasibility of fitting new/future applications/services into non-harmonized spectrum of the digital dividend (namely the so-called "white spaces" between allotments. Report C from CEPT to the European Commission in response to the Mandate on: Technical considerations regarding harmonization options for the Digital Dividend. 1 July 2008.
- [12] Mitola, J. and Maguire, G., (1999) Cognitive radio: making software radios more personal. *IEEE Personal Communications Magazine*, Volume 6 Issue 4. pp. 13 to 18.
- [13] Mitola, J. (2000) Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio. Dissertation submitted in partial fulfillment of the degree of Doctor of Technology. Royal Institute of Technology (KTH) - Sweden. Teleinformatics. ISSN 1403 - 5286.
- [14] Cordeiro, C., Kiran Challapali, and Dagnachew Birru, Sai Shankar (2006) N. IEEE 802.22: An Introduction to the First Wireless Standard based on Cognitive Radios *Journal of Communications*, Vol. 1, N° 1.
- [15] Castro Lechtaler, A.; Foti, A.; Arroyo Arzubi, A.; García Guibout, J.; Carmona, F.; Fusario, R. y Oliveros, A. Using White Spaces: A solution for frequency spectrum overloading. XI Workshop of Architecture, Networks and Operating Systems (WARSO). *Proceedings of the 22th Argentinean Congress on Computer Science*. ISBN 978-987-733-072-4. pp. 903 to 913. San Luis. October. 2016.
- [16] Castro Lechtaler, A; Arroyo Arzubi, A; Foti, A.; Fusario, R.; García Guibout, J.; Oliveros, A. y Carmona, F. Using White Spaces: A solution for frequency spectrum overloading. *Computer Science & Technology Series. XXII Argentine Congress of Computer Science. Selected Papers*. ISBN: 978-987-4127-28-0. Pg. 235 to 246. EDULP. 2017.

Análisis de Protocolos de Comunicaciones para Internet de las Cosas

Mg. Jorge Eterovic; Esp. Marcelo Cipriano; Santiago Nicolet

Instituto de Investigación en Ciencia y Tecnología
Dirección de Investigación Vicerrectorado de Investigación y Desarrollo.
Universidad del Salvador.
Lavalle 1854 – C1051AAB -Ciudad Autónoma de Buenos Aires - Argentina

jorge.eterovic@gmail.com; cipriano1.618@gmail.com; santiago.nicolet@usal.edu.ar

RESUMEN

Internet de las cosas (en inglés, Internet of Things, abreviado: IoT) [1,2] es un concepto que se refiere a la interconexión digital de cosas u objetos en Internet [3]. Según la consultora Gartner [4], en 2020 habrá en el mundo aproximadamente 26 mil millones de dispositivos con un sistema de conexión a Internet de las cosas.

Entre las tecnologías de comunicaciones más usadas en IoT se encuentran: RFID - Radio Frequency Identification, NFC - Near Field Communication y WSN - Wireless Sensor Networks.

Para el usuario de IoT, estas tecnologías resultan ser “transparentes”. Es decir que se ignora su existencia o se tiene una visión parcial o incompleta de las mismas. Esta “transparencia” también incluye a cuáles son los protocolos adecuados para cada tipo de aplicación y las técnicas de protección y seguridad de las comunicaciones y del transporte y almacenamiento de datos confidenciales y/o sensibles, en los sistemas que así lo requieren.

Esta investigación se centrará en encontrar los indicadores que permitan identificar la mejor solución de comunicaciones en Internet de las Cosas, y que tenga la capacidad de incorporar soluciones de seguridad, tales como Criptografía Ligera o Liviana [5], para garantizar la privacidad y la protección de los datos personales [6].

Palabras Clave:

Internet de las Cosas, Protocolos de Comunicaciones en IoT, Seguridad en IoT.

CONTEXTO

El Vicerrectorado de Investigación y Desarrollo (VRID), perteneciente a la Universidad del Salvador (USAL), dicta las políticas referidas a la investigación, concibiéndola como un servicio a la comunidad, entendiendo que los nuevos conocimientos son la base de los cambios sociales y productivos. Con el impulso de las propias Unidades Académicas, se han venido desarrollando acciones conducentes a concretar proyectos de investigación uni/multidisciplinarios, asociándolos a la docencia de grado y postgrado y vinculando este accionar, para potenciarlo, con otras instituciones académicas del ámbito nacional e internacional.

La Dirección de Investigación, dependiente del VRID, brinda soporte a las distintas Unidades de Investigación y a sus investigadores para el desarrollo de Proyectos y Programas de Investigación, nacionales e internacionales, como así también, apoyo y orientación de recursos para la investigación.

A ella pertenece el Instituto de Investigación en Ciencia y Tecnología (RR 576/12) en el cual se enmarca este proyecto, con una duración de 2 años (2017-2018).

1. INTRODUCCIÓN.

En 1999, Kevin Ashton, miembro fundador del Laboratorio de Investigación Auto-ID Center del MIT [7], hoy llamado Auto-ID Labs, fue el primero que acuñó y usó el término Internet of things (IoT), donde se realizaban investigaciones en el campo de la identificación por radiofrecuencia (RFID) y tecnologías de sensores.

En IoT cada objeto, ya sea virtual o físico, es transmisible, direccionable y accesible a través de Internet. Cada objeto tiene su propia identificación y tiene la capacidad de detectar, procesar y comunicarse [8].

La naturaleza omnipresente de los objetos en IoT hace que los datos que se recopilan y transmiten para uso público y privado sean muy importantes y se debe garantizar la seguridad de los mismos. La integridad y la confidencialidad de los datos transmitidos, así como la autenticación de los objetos son los aspectos clave de la seguridad y de la privacidad en IoT.

En la Fig. 1 se muestra el Hype Cycle elaborado por la consultora Gartner [9] del estado de las tecnologías emergentes. El Hype Cycle es un gráfico que muestra el estado de madurez de la adopción y de la aplicación de una tecnología. IoT, marcada con una línea roja, se encuentra en la fase ascendente de la curva, conocida como “tecnologías disparadoras de innovación”.

La seguridad y la privacidad son un tema extenso que cubre toda la pila de protocolos de comunicaciones. Los principales problemas de seguridad en IoT incluyen Autenticación, Identificación y heterogeneidad del dispositivo. Dado que cada dispositivo tiene su propia identificación, será muy difícil identificar miles de millones de dispositivos.

Autenticar cada dispositivo puede ser un trabajo tedioso. Una de las principales preocupaciones de seguridad es la heterogeneidad de dispositivos, que impide aplicar una única solución de seguridad uniforme en todos los casos.

Cada dispositivo tiene diferentes requerimientos de seguridad. La heterogeneidad del dispositivo también puede causar problemas en otros aspectos.



Fig. 1.

La historia de la seguridad de los dispositivos se inicia con las conocidas “etiquetas antirrobo” que se adhieren a libros, prendas y demás objetos en librerías y shoppings. Luego aparecieron otros objetos, como las llaves “codificadas” de vehículos, los “tags” para abonar peajes y tarjetas para el pago electrónico de pasajes en transporte público (tarjetas Monedero, SUBE, etc.). Pero menos conocidos por su reciente aparición y no tan masiva difusión como son los pasaportes, licencias de conducir, documentos de Identidad y hasta incluso minúsculos chips subcutáneos, entre otros dispositivos y sistemas a implementar.

Obviamente existe una gran diferencia entre la Internet convencional e IoT [10]. Las principales características de IoT son: procesamiento lento, memoria limitada y baja potencia. Las redes de IoT se conocen generalmente como redes con pérdidas y de baja potencia (LLN - Low power and Lossy Networks) dado que son susceptibles a que sufran una gran pérdida de datos.

Las principales tecnologías de comunicaciones utilizada en IoT [11] son:

- RFID: Radio Frequency Identification [12]
- WSN: Wireless Sensor Network [13]
- NFC: Near Field Communication [14]
- WiFi: estándar IEEE 802.11n
- Bluetooth

- 4G: la red de telefonía móvil
- LTE - Long Term Evolution
- ZigBee: estándar IEEE 802.15.4
- IEEE 802.11ah
- Z-Wave
- Sigfox
- LoRaWAN

En general, podemos decir que están siendo utilizadas diferentes tecnologías de comunicaciones, dependiendo de la aplicación y sus requerimientos de alcance, volumen de datos, seguridad, consumo de energía, vida útil de la batería, etc.

Las redes de comunicaciones han ido evolucionando hacia el sector del IoT que, aunque actualmente no compite con el sector de la telefonía móvil a nivel comercial, pero ha despertado el interés y la inversión de numerosas empresas en este sector [15].

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO.

Se realizará un relevamiento, estudio y análisis exhaustivo de los principales protocolos de comunicaciones que podrían ser usados en IoT.

Se analizarán los protocolos para determinar el grado de exposición en los aspectos de privacidad, protección de datos personales y seguridad en las comunicaciones.

Se definirán indicadores para evaluar comportamientos y permitir comparaciones utilizando las experiencias publicadas en trabajos internacionales.

Se volcarán los resultados obtenidos en una tabla comparativa y en gráficos de usabilidad de los protocolos estudiados.

Finalmente se redactará un informe final con los resultados obtenidos.

3. RESULTADOS OBTENIDOS / ESPERADOS.

El objetivo de este proyecto es realizar un análisis comparativo de los protocolos de comunicaciones, de acuerdo con, por ejemplo,

indicadores de alcance, velocidad de transmisión de los datos y consumo de energía, para darle seguridad usando Criptografía Ligera o Liviana.

El uso de estos indicadores nos permitirá evaluar comportamientos y permitir comparaciones a fin de poder seleccionar el mejor protocolo de comunicaciones según las necesidades del usuario.

Así se espera que se puede determinar de manera rápida que protocolo de comunicaciones sería recomendable para transmitir una dada cantidad de datos a una determinada velocidad y a una cierta distancia.

También se identificará el nivel de adopción de los distintos protocolos. Tendremos así la posibilidad de establecer cuáles ya ha sido adoptados por la industria, los nuevos estándares publicados en las normas, pero aún no implementados y los desarrollos de próxima evaluación.

Finalmente se redactará un informe final y se presentarán los resultados obtenidos de esta investigación en diferentes congresos, para su difusión y como un aporte al conocimiento de la comunidad científica.

4. FORMACIÓN DE RECURSOS HUMANOS.

El equipo de investigadores pertenece al cuerpo docente de Tecnologías Aplicadas en la Facultad de Ingeniería, el área de la Seguridad Informática, de la Universidad del Salvador.

A fines del año 2017 se han incorporado el Ing. Santiago Nicolet, docente de las carreras de Ingeniería en Informática y de la Licenciatura en Sistemas de Información y el alumno Damián Rodríguez, como colaboradores en el equipo de investigación. Se espera que en breve se incorporen más alumnos.

5. BIBLIOGRAFÍA.

- [1] Internet of Things. <http://www.cisco.com/web/solutions/trends/iot/overview.html>. Ultima vez consultada: febrero de 2018.
- [2] Internet de las cosas Cómo la próxima evolución de Internet lo cambia todo <http://www.cisco.com/web/LA/soluciones/execute/assets/pdf/internet-of-things-iot-ibsg.pdf>. Ultima vez consultada: febrero de 2018.
- [3] Conner, Margery; Sensors empower the "Internet of Things"; Issue 10; pp. 32-38. Mayo de 2010; ISSN 0012-7515
- [4] Gartner Says the Internet of Things Installed Base Will Grow to 26 Billion Units By 2020. <https://www.gartner.com/newsroom/id/2636073>. Ultima vez consultada: febrero de 2018.
- [5] ISO/IEC 29192. Information technology - Security techniques - Lightweight Cryptography. 2012. <https://www.iso.org>.
- [6] Román R., Nájera P., López J. "Los Desafíos De Seguridad En La Internet De Los Objetos" University of Malaga, España. 2010.
- [7] Kevin Ashton; That 'Internet of Things' Thing; <http://www.rfidjournal.com/articles/view?4986>. Ultima vez consultada: febrero de 2018.
- [8] Advancing the IoT for Global Commerce <https://autoidlabs.org/>. Ultima vez consultada: febrero de 2018.
- [9] Gartner's Hype Cycle Special Report for 2017, Gartner Inc. <https://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017/>. Ultima vez consultada: febrero de 2018.
- [10] Differences between the IoT and Traditional Internet by Dr. Opher <https://www.rtinsights.com/differences-between-theiot-and-traditional-internet/> Ultima vez consultada: febrero de 2018.
- [11] Laeeq, Kashif, and Jawwad A. Shamsi, "A Study of Security Issues, Vulnerabilities and Challenges in Internet of Things," Securing Cyber-Physical Systems, p. 221, 2015.
- [12] Radio frequency identification ready to deliver Armed forces communications and electronics association 2005. <http://www.afcea.org>
- [13] <http://www.lanacion.com.ar/1892969-club-tigre-chips-bajo-la-piel-una-tecnologia-de-identificacion-practica-o-invasiva>. Ultima vez consultada: febrero de 2018.
- [14] Masanobu Katagi; Shiho Moriai, Lightweight Cryptography for the Internet of Things; Sony Corporation; 2016.
- [15] Tecnologías de Comunicaciones para IoT. Link: <https://www.efor.es/servicios/internet-de-las-cosas-iot.html>. Ultima vez consultada: febrero de 2018.

Hub Of Things: Concentrador para el Internet de las Cosas

Ricardo Brea¹, Daniel Skrie¹, Marisa Panizzi¹

¹Escuela de Sistemas. Universidad Argentina John F. Kennedy.
Bartolomé Mitre 1411, Ciudad Autónoma de Buenos Aires (C1037ABA), Tel: 5236-1200
brea.ricardo@gmail.com; dskrie@yahoo.com.ar; marisapanizzi@outlook.com

RESUMEN

En la actualidad, cada vez encontramos más dispositivos electrónicos conectados a internet, monitoreados y controlados en forma remota. La diversidad tecnológica y la cantidad de dispositivos dificulta la integración de los mismos para su control, monitoreo e interacción. Las plataformas basadas en microcontroladores o de procesamiento reducido como *Arduino*, no brindan una conexión con niveles aceptables de seguridad. La privacidad constituye una dificultad, ya que los usuarios desconocen si los distintos proveedores de soluciones *IoT* utilizan sus datos o los venden a terceros. La latencia también es un problema. Muchos de los proveedores de soluciones en la nube no tienen servidores locales, lo que degrada el tiempo de reacción ante determinado evento. Se propone el *Hub Of Things* o *HoT* como una solución que integre localmente o en la nube, el control y monitoreo de los dispositivos. Proveerá además una interface de control, segura y homogénea, tanto gráfica como de programación. Será escalable porque contemplará un método que amplía la variedad de dispositivos a integrar y monitorear, posibilitando su interacción con otros sistemas. El *HoT* intentará solucionar los problemas de la diversidad tecnológica, por medio de una interface homogénea y segura. Sirviendo de mediador entre los dispositivos y el usuario.

Palabras clave: Internet de las Cosas, *Fog computing*, Seguridad, *MQTT*.

CONTEXTO

Este trabajo se desarrolla en la Escuela de Sistemas de la Universidad John F. Kennedy, en el marco de la asignatura Taller de Trabajo Final Integrador de la carrera Licenciatura en Sistemas

1. INTRODUCCION

El término “Internet de las Cosas” o *IoT* por sus siglas en inglés, es un concepto acuñado en 1999 por Kevin Ashton investigador de MIT [1]. Originalmente los datos disponibles en internet eran ingresados o generados por humanos, con la incorporación de sensores y conectividad más accesible. Los dispositivos o “cosas” son capaces de generar datos y pueden ser muy diversos: teléfonos inteligentes, bandas inteligentes (*SmartBands*), estaciones meteorológicas, automóviles, etc.

El *IoT* genera grandes posibilidades de crecimiento, no solo en el ámbito hogareño, también en el industrial conocido como Industria 4.0, o bien en la planificación urbana bajo el concepto de ciudad inteligente o *Smart Cities*.

En la Tabla 1, se representan las distintas proyecciones para el crecimiento de *IoT*, que promedian 25,7 billones de dispositivos para 2020.

Tabla 1 Pronósticos de crecimiento del IoT

Fuente	Billones de Dispositivos
Cisco [2]	26.3
Ericsson [3]	28
Gartner [4]	20.8
Goldman Sachs [5]	28

Con un pronóstico de crecimiento tan optimista es importante detectar los problemas que actualmente presenta el IoT. *Acquity Group* (División de *Accenture* dedicada a la estrategia digital y marketing) realizó en el año 2014 una encuesta a 2000 consumidores de los Estados Unidos, que conociendo los dispositivos *IoT*, no los adquirirían por los siguientes motivos;

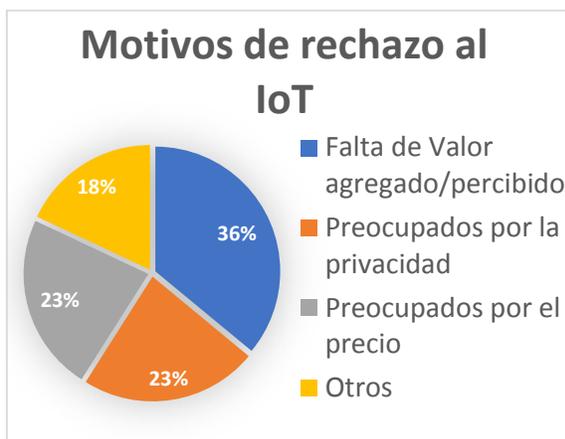


Figura 1. Motivos de rechazo del IoT [6]

La falta de valor agregado es percibida, entre otros motivos, a partir de las dificultades de integración con soluciones similares o existentes.

La preocupación por la privacidad es un elemento inherente a las soluciones en la nube. Los proveedores de estos servicios mitigan el problema con un contrato de privacidad.

La preocupación por el precio responde a varios elementos, como por ejemplo el desconocimiento de las tecnologías y la incertidumbre en los costos de los servicios que formarán parte del costo fijo.

Hay aspectos técnicos que impiden la implementación del *IoT*, como la latencia, el tiempo de respuesta de la solución. Los grandes proveedores de servicios en la nube (Microsoft, Amazon, Oracle e IBM) tienen sus servidores fuera del país y en soluciones industriales, el tiempo de reacción es un factor importante. Para reducir la latencia inherente de las soluciones en la nube surge el

concepto de *Fog Computing* [7] o *Edge Computing* que se refiere a:

- Analizar los datos más sensibles en el *Edge*, el borde, donde se generan los datos, sin necesidad de enviar grandes volúmenes de datos a la red.
- Reaccionar a la información generada en el rango de milisegundos.
- Enviar la información a la nube para su análisis y almacenamiento a largo plazo.

El objetivo del *Hub Of Things* es proveer una plataforma segura, homogénea, extensible y abierta, que opera en el *Fog Computing* y ser utilizada como base en nuevas soluciones *IoT* para el control y monitoreo de dispositivos conectados a internet. El *Hub Of Things* plantea una solución de bajo costo frente a los motivos de rechazo al IoT.

En la figura 2, se presenta que el *Hub Of Things* como un *IoT Gateway*, una solución de software y hardware que serviría de intermediario entre los dispositivos y otros sistemas, ya sea que estén implementados localmente (*On-Premise*) o en la Nube. Se incrementaría la seguridad y estabilidad, reduciéndose la latencia. Se considera que al brindar una capa de abstracción, se reduce el acoplamiento con servicios *IoT* en la nube y permite adaptar otros protocolos de comunicación como *LoRa*, *BLE*, *Modbus*, etc. a un único protocolo, que en este caso es *MQTT*.

MQTT (*Message Queue Telemetry Transport*), es un protocolo usado para la comunicación *machine-to-machine* (*M2M*) en *IoT*. Está orientado a la comunicación de sensores, debido a que minimiza la información adicional a cada mensaje que envía y puede ser utilizado en la mayoría de los dispositivos embebidos y de pocos recursos. *MQTT* sigue una topología de estrella, con un nodo central que hace de servidor o *bróker* y distribuye los mensajes utilizando el patrón *publiser-subscriber* [8].

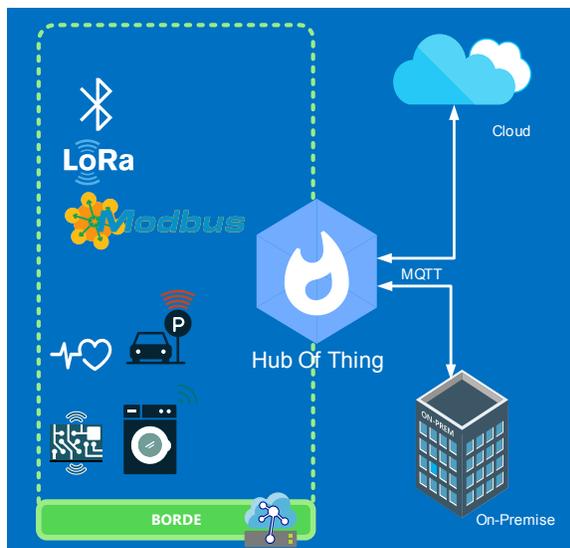


Figura 2 - Arquitectura de alto nivel - HoT

Internamente el *Hub Of Things* contemplará la siguiente arquitectura:

Capa de adaptación: El objetivo de esta capa será proveer una interface de programación, *API*, que permitirá adaptar cualquier protocolo al sistema de mensajes similar a *MQTT*, es decir en tuplas $\langle \langle \text{TOPIC}, \text{PAYLOAD} \rangle \rangle$.

MQTT Bróker: Se implementará con una librería abierta denominada Mosca que implementa el protocolo MQTT sobre TCP y *WebSockets* facilitando la comunicación con clientes JavaScript.

Interface HTML: La capa de presentación se encontrará programada en *HTML5* y *JavaScript* utilizando el *framework Angular* para aplicaciones de una sola página o “*SPA*”. El diseño de la página seguirá la directiva *Material Desing* de Google [9], que se asemeja a los controles del sistema operativo *Android*. Dichas directivas existen para diseñar interfaces de usuario intuitivas, homogéneas y fáciles de utilizar. Otras características de la interface HTML son; Interface *Responsive*, adaptable a varios tipos de dispositivos y gráficos *SVG*, vectoriales que consumen menos espacio y son visibles en cualquier resolución [10].

2. LINEAS DE INVESTIGACION Y DESARROLLO

El *Hub Of Things* es una herramienta que constituye un punto de inicio para otras soluciones *IoT*, vinculadas a problemas particulares o como integradora de soluciones que implementan protocolos no TCP/IP. Las líneas de investigación se verán orientadas a la recopilación de datos y la integración de sistemas. Si bien cada línea puede enfocarse como un trabajo independiente, son mencionadas por que utilizarán el *Hub Of Things* como elemento central.

A saber:

- Integración de sistemas de procesamiento de imágenes como el índice de vegetación de diferencia normalizada o NVDI [11].
- Medición de consumo eléctrico hogareño.
- Monitoreo de variables ambientales urbanas como la concentración de CO_2 o el nivel de ruido ambiente.
- Integración de sistemas inalámbricos *IoT* de largo alcance con *LoRa*.
- Integración de sistema cableados *Modbus*.

3. RESULTADOS OBTENIDOS / ESPERADOS

El objetivo principal de este trabajo es proveer una solución segura, homogénea, extensible y abierta para el control y monitoreo de dispositivos conectados a internet.

Los objetivos de este trabajo son:

- Proveer una *API* que permita la integración de otros dispositivos y protocolos al sistema.
- Brindar flexibilidad a las estructuras de datos para almacenar la configuración de los dispositivos integrados, de forma tal que se pueda incorporar una gran variedad de datos, dada la diversidad de opciones

- Implementar un *Bróker MQTT* integrado con capacidad de conectarse en modo puente a otro *Bróker* central.
- Implementar autenticación en todas las interfaces.
- Desarrollar interface web de usuario intuitiva y flexible tanto a nivel usuario como de programación.
- Implementar seguridad a nivel de transporte. Es decir cifrar las comunicaciones siempre que sea posible desde y hacia los clientes.
- Utilizar componentes propios o de terceros con el único requisito de que sean de código fuente abierto. Para mantener el costo de licencias en cero.
- Definir una arquitectura simple para que pueda ser utilizado tanto en *PC* o *SBC* (computadoras en una placa) como *Raspberry Pi*.

La Pregunta Problema:

¿La incorporación de una solución abierta, homogénea, segura y de bajo costo incrementara la adopción de soluciones IoT en los mercados hogareños (*Smart home*), corporativos (*Smart buildings*), urbano (*Smart Cities*) e industriales (Industria 4.0)?

4. FORMACION DE RECURSOS HUMANOS

En este trabajo que se desarrolla en la Escuela de Sistemas de la Universidad John F. Kennedy, participan un alumno avanzado de la carrera Licenciatura en Sistemas que se encuentra desarrollando su trabajo de fin de carrera, el docente a cargo de la asignatura y la Directora de la Escuela.

5. BIBLIOGRAFIA

- [1] K. Ashton, «That 'Internet of Things' Thing,» 22 JUN 22. [En línea]. Available: <http://www.rfidjournal.com/articles/view?4986>. [Último acceso: FEB 2016].
- [2] Cisco, «Cisco Visual Networking Index Predicts Near-Tripling of IP Traffic by 2020,» 07 JUN 2016. [En línea]. Available: <https://newsroom.cisco.com/press-release-content?type=press-release&articleId=1771211>. [Último acceso: FEB 2018].
- [3] Ericsson, «Ericsson Mobility Report November 2015,» NOV 2015. [En línea]. Available: <https://www.ericsson.com/assets/local/news/2016/03/ericsson-mobility-report-nov-2015.pdf>. [Último acceso: FEB 2018].
- [4] Gartner, «Gartner Says 6.4 Billion Connected "Things" Will Be in Use in 2016,» 10 NOV 2015. [En línea]. Available: <https://www.gartner.com/newsroom/id/3165317>. [Último acceso: FEB 2018].
- [5] S. Jankows, J. Covello, H. Bellini, J. Ritchie y D. Costa, «The Internet of Things: Making sense of the next mega-trend,» 3 SEP 2014. [En línea]. Available: <http://www.goldmansachs.com/our-thinking/outlook/internet-of-things/iot-report.pdf>. [Último acceso: FEB 2018].
- [6] Accenture: Acquity Group, «IoT: The Future of Consumer Adoption,» 2014. [En línea]: https://www.accenture.com/t20150624_T211456_w_/us-en/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Technology_9/Accenture-Internet-Things.pdf. [Último acceso: NOV 2016].
- [7] Cisco, «Fog Computing and the Internet of Things,» 2015. [En línea]: https://www.cisco.com/c/dam/en_us/solutions/trends/iot/docs/computing-overview.pdf. [Último acceso: FEB 2016].
- [8] OASIS, «MQTT Version 3.1.1,» OCT 29 2014. [En línea]. Available: <http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html>. [Último acceso: ENE 2016].

- [9] Google, «Material Design - Guidelines,» 2014. [En línea]. Available: <https://material.io/guidelines/>. [Último acceso: NOV 2016].
- [10] D. Strazzullo, D. D. Dailey and J. Frost, *Building Web Applications with SVG*, Sebastopol, California: O'Reilly Media, Inc., 2012.
- [11] J. Weier y D. Herring, «NASA-Earthobservatory-Measuring Vegetation (NDVI & EVI),» 30 AUG 2000. [En línea]. Available: https://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_1.php. [Último acceso: JUN 2017].

“Análisis de Eficiencia y Tolerancia a Fallos en Arquitecturas Multiprocesador para Aplicaciones de Procesamiento de Datos”

Jorge R. Osio^{1,2}, Diego Montezanti^{1,4}, Eduardo Kunysz¹, Diego Encinas^{1,4} Daniel Martin Morales^{1,3}

¹Instituto de Ingeniería y Agronomía - UNAJ

²UIDET CeTAD –Fac. de Ingeniería - UNLP

³Laboratorio LINES - Universidad Tecnológica Nacional FRLP

⁴Instituto de Investigación en Informática LIDI - Fac. de Informática – UNLP

{josio, dmontezanti, ekunysz, martin.morales}@unaj.edu.ar

Resumen

Dentro de la línea de investigación que se está desarrollando, existen dos enfoques diferentes. Por un lado se está trabajando sobre la implementación de algoritmos de procesamiento de imágenes sobre dispositivos reconfigurables. El objetivo es utilizar una combinación de diferentes técnicas de concurrencia y paralelismo para tener en cuenta aspectos comunes de dichos algoritmos, y así mejorar la eficiencia en el procesamiento de imágenes médicas. Por otra parte, debido a que el procesamiento en paralelo requiere de la implementación de sistemas de múltiples procesadores en dispositivos reconfigurables, resulta útil incorporar al proyecto el desarrollo de metodologías que permitan tolerar fallos transitorios que son característicos de las arquitecturas multicore, y que afectan especialmente la ejecución de aplicaciones paralelas de cómputo intensivo.

Palabras clave: *arquitecturas paralelas, procesamiento de imágenes, tolerancia a fallos, sistemas multicores, dispositivos reconfigurables.*

Contexto

La línea de Investigación descrita en este trabajo forma parte del Proyecto de Investigación Científico-Tecnológico “Tecnologías de la información y las comunicaciones mediante IoT para la solución de problemas en el medio socio productivo”, que se desarrolla en la Universidad Nacional Arturo Jauretche (UNAJ) y es continuación del proyecto presentado en la edición anterior de WICC [7].

El proyecto cuenta además con financiamiento en el marco del programa “Universidad, Diseño y Desarrollo Productivo” del Ministerio de Educación a través del proyecto “Sistema de eficiencia energética” y está a la espera de su Aprobación para obtener financiamiento de la convocatoria “UNAJ Investiga 2017”.

Parte de las líneas de investigación desarrolladas se encuentran enmarcadas en los convenios de colaboración en Actividades de Investigación firmados por la UNAJ con la UIDET-CeTAD.

1. Introducción

En los últimos años, se ha buscado expandir el concepto del procesamiento paralelo basado en computadoras con multicore, llevándolo hacia la utilización de plataformas más específicas. Para obtener una mayor eficiencia en el procesamiento, los fabricantes de computadoras de altas prestaciones han introducido unidades basadas en FPGAs (*Field-Programmable Gate Array*) en su diseño como soporte para el cómputo [1-3].

Si bien el estudio de sistemas paralelos con multiprocesadores es un campo bien desarrollado, la utilización de múltiples cores en sistemas reconfigurables es un terreno que tiene múltiples posibilidades de exploración [4]. En el presente proyecto se exploran mejoras en la implementación de los algoritmos mediante procesamiento paralelo mediante procesadores microblaze y concurrente en VHDL (lenguaje de descripción de hardware) y la tolerancia a fallos en entornos multicore.

Plataformas FPGAs para procesamiento paralelo

La implementación del paralelismo en plataformas FPGAs consiste en el uso de procesadores embebidos para ejecutar aplicaciones y en la utilización de las características que provee la lógica programable para manejar las porciones de código que se ejecutan concurrentemente [4].

La facilidad de implementar procesadores embebidos en forma rápida, junto con la posibilidad de proveer concurrencia mediante la programación en hardware, permiten combinar las FPGAs con el paralelismo obtenido

mediante sistemas multicore para alcanzar una alta eficiencia, que es un gran desafío en la búsqueda de optimizar el procesamiento de imágenes médicas. La mejora se aplica al conjunto de algoritmos basados en operadores de ventana, donde una porción de los algoritmos es común a todos posibilitando la concurrencia en VHDL, mientras que la otra parte del algoritmo se implementa en software y se ejecuta en paralelo dentro del sistema multicore.

Con las mejoras constantes que aporta la evolución de la tecnología sobre las FPGAs pueden lograrse diseños de gran magnitud, a tal punto que la tendencia actual es implementar microprocesadores de propósito general, conjuntamente con todo el hardware de propósito específico que requiere la aplicación, dentro de una FPGA.

Tolerancia a Fallos en sistemas multicore

El aumento en la escala de integración, con el objetivo de mejorar las prestaciones en los procesadores actuales, sumado al crecimiento de los sistemas de cómputo, han producido que la confiabilidad se haya vuelto un aspecto relevante. En particular, la creciente vulnerabilidad a los fallos transitorios se ha vuelto crítica, a causa de la capacidad de estos fallos de corromper los resultados de las aplicaciones.

El impacto de los fallos transitorios se vuelve más notorio en el contexto del HPC, donde el escenario típicamente consiste en una supercomputadora con cientos o miles de procesadores que cooperan para ejecutar aplicaciones paralelas, debido a que el Tiempo Medio Entre Fallos (MTBF) del sistema disminuye cuanto mayor es la cantidad de procesadores. La incidencia de los fallos transitorios es aún mayor en el caso de

aplicaciones intensivas en cómputo y de gran duración, que manejan elevados volúmenes de datos [11].

El alto costo (en términos temporales y de utilización de recursos) que implica volver a lanzar una ejecución desde el comienzo, en caso de que un fallo transitorio produzca la finalización de la aplicación con resultados incorrectos, justifica la necesidad de desarrollar estrategias específicas para mejorar la confiabilidad y robustez en sistemas de HPC. En particular, resulta crucial poder detectar los fallos llamados silenciosos, que alteran los resultados de las aplicaciones pero que no son interceptados por el sistema operativo ni ninguna otra capa de software del sistema, por lo que no causan la finalización abrupta de la ejecución.

Antecedentes del Grupo de Trabajo

Los antecedentes del grupo de trabajo parten de la investigación acerca de procesamiento de imágenes sobre diferentes arquitecturas paralelas [4-6] y de la investigación de la Tolerancia a fallos en sistemas multicore [8-9].

Las líneas de investigación se enmarcan fundamentalmente en la búsqueda de técnicas híbridas de procesamiento paralelo [10], sobre diferentes arquitecturas, que optimicen las prestaciones en sistemas de adquisición y procesamiento de imágenes médicas. En particular, resulta de interés su incorporación y aplicación en el Hospital “El Cruce” de Florencio Varela. Adicionalmente, se pretenden encausar temas de estudios de posgrado actuales y futuros de docentes y profesionales surgidos de la UNAJ dentro de estas líneas de investigación. El director del proyecto participa en proyectos de investigación desde el año 2005 en la Facultad de Ingeniería de la UNLP como

colaborador, y a partir de 2010 como investigador. Además dirige proyectos de investigación desde el año 2015 en la UNAJ. Actualmente, el grupo de investigación enfoca su trabajo en los temas: Procesamiento Digital de Imágenes Médicas sobre plataformas FPGA, Procesamiento de Imágenes en arquitecturas FPGA multiprocesador y Tolerancia a Fallos en Sistemas de Cómputo de Altas Prestaciones, pertenecientes a estudios de Maestrías y Doctorados que se realizan en el marco de acuerdos de colaboración entre el Laboratorio CeTAD, el Instituto III-LIDI y la UNAJ, a través de los respectivos proyectos de investigación.

Como resultado del trabajo realizado recientemente en la UNAJ se han publicado artículos en diferentes congresos nacionales e internacionales de Ingeniería Informática y Ciencias de la Computación [3-8]

2. Líneas de Investigación y Desarrollo

El grupo de investigación que se ha constituido recientemente en la UNAJ es multidisciplinario, y sus miembros cuentan con experiencia en sistemas multiprocesador, sistemas embebidos, procesamiento de imágenes y tolerancia a fallos.

Temas de Estudio e Investigación

- Implementación de un sistema multiprocesador en Dispositivos Lógicos Programables (FPGAs).
- Análisis y determinación de la eficiencia obtenida en el procesamiento de imágenes mediante la aplicación de técnicas de concurrencia y paralelismo.

- Implementación de una metodología distribuida basada en software para proveer tolerancia a fallos transitorios, específicamente en entornos de clusters de multicores donde se ejecutan aplicaciones paralelas científicas de paso de mensajes.

3. Resultados Obtenidos/Esperados

Investigación experimental

Hasta el momento se han obtenido los siguientes resultados en relación a los objetivos principales:

- En cuando al procesamiento de imágenes sobre dispositivos FPGAs, se han implementados varios algoritmos en sistemas multicore con buenos resultados en el procesamiento paralelo. En donde uno de cores se encarga de recibir y distribuir los datos para el posterior procesamiento.
- En cuanto a la tolerancia a fallos transitorios, se ha implementado una estrategia de detección basada en el monitoreo de las comunicaciones, y se ha diseñado un mecanismo de recuperación automático basado en el almacenamiento de un conjunto de *checkpoints* distribuidos de capa de sistema.

Para el año en curso, se esperan alcanzar importantes resultados en el área de concurrencia y cómputo paralelo, posibilitados por la utilización de sistemas basados en FPGAs. Con esta arquitectura, constituida por varios procesadores implementados en una misma FPGA, se espera lograr mayor eficiencia mediante la combinación de técnicas de paralelismo y la concurrencia lograda en VHDL. Para las pruebas se

emplearán algoritmos de procesamiento de imágenes basados en operadores de ventana, debido a que una parte del procesamiento es común para todos, por lo que puede implementarse de forma concurrente mediante un co-procesador; mientras que el resto puede paralelizarse entre todos los procesadores. La implementación de diferentes algoritmos permitirá medir la performance en la ejecución alcanzada con el sistema multiprocesador, y sacar conclusiones a partir de las características de cada algoritmo.

En cuanto a la tolerancia a fallos, se ha propuesto una metodología distribuida basada en replicación de software, diseñada específicamente para aplicaciones paralelas científicas de paso de mensajes, capaz de protegerlas de fallos transitorios que producirían ejecuciones incorrectas [9]. Bajo la premisa de que, en este tipo de aplicaciones, la mayor parte de los datos relevantes para el resultado son transmitidos entre procesos, la estrategia de detección se basa en validar los contenidos de los mensajes que se van a enviar y comparar los resultados finales, obteniendo un compromiso entre un alto nivel de cobertura frente a fallos y la introducción de un bajo *overhead* temporal y sobrecarga de operaciones, debido a que no se realiza trabajo para detectar fallos que no afectan a los resultados. Por otra parte, previene la propagación de la corrupción de datos a otros procesos, manteniéndola aislada en el contexto del proceso que ha resultado directamente afectado. De esta manera, se mejora la confiabilidad del sistema y se disminuye el tiempo luego del cual se puede relanzar la aplicación, lo cual es especialmente útil en ejecuciones prolongadas que conllevan un alto coste. Por otra parte, aprovecha el beneficio del

sistema la redundancia intrínseca de hardware de las arquitecturas multicore.

Para tolerar completamente los fallos transitorios, la propuesta se basa en restaurar el sistema a un estado consistente previo a su ocurrencia. Para ello, se debe integrar la estrategia de detección con mecanismos de protección para fallos permanentes. Se están estudiando las opciones de incorporar recuperación basada en múltiples *checkpoints* coordinados en capa de sistema (que proporcionan cobertura en el caso de que un *checkpoint* resulte afectado por un fallo, haciendo imposible la recuperación), o la utilización de un único *checkpoint* no coordinado de capa de aplicación, que puede ser verificado para asegurar que no contiene errores.

El uso de estas estrategias posibilitaría prescindir de la utilización de redundancia triple con votación para detectar y recuperar de fallos transitorios. Además, como estos fallos no requieren reconfiguración del sistema, la recuperación puede realizarse mediante re-ejecución en el mismo core en el que ocurrió el fallo.

4. Formación de Recursos Humanos

Dentro de la temática de la línea de I+D, todos los miembros del proyecto participan en el dictado de asignaturas de la carrera de Ingeniería Informática de la UNAJ.

En este proyecto existe cooperación a nivel nacional. Hay dos investigadores realizando Doctorados y dos realizando Maestrías en temas relacionados con simulación de sistemas multiprocesador, sistemas embebidos, software embebido, sistemas multicore y tolerancia a fallos en HPC.

Adicionalmente, se cuenta con la colaboración de estudiantes avanzados.

5. Bibliografía

- [1] O. Mencer, K. Tsoi, S. Cramer, T. Todman, W. Luk, Ming Wong and P. Leong, "CUBE: a 512-FPGA Cluster", Dept. of Computing, Imperial College London, Dept. of Computer Science and Engineering The Chinese University of Hong Kong. (2009)
- [2] Keith Underwood, "FPGAs vs. CPUs: Trends in Peak Floating-Point Performance", Sandia National Laboratories. (2011)
- [3] D. Encinas, E. Kunysz, Szymanowski, Morales M., "Performance de arquitecturas multiprocesador: técnicas de simulación y plataformas reconfigurables", Instituto de Ingeniería, UNAJ, (2014)
- [4] J. Osio, J. Salvatore, E. Kunysz, V. Guarepi, M. Morales, "Análisis de Eficiencia en Arquitecturas Multiprocesador para Aplicaciones de Transmisión y Procesamiento de Datos", UNER, Ciudad de Concordia, WICC 2016.
- [5] J. Osio, D. Montezanti, M. Morales, "Análisis de Eficiencia en Sistemas Paralelos", Ushuaia, Tierra del Fuego, WICC 2014
- [6] E. Kunysz, J. Rapallini, J. Osio, "Sistema de cómputo reconfigurable de alta performance (Proyecto HPRC)", 3ras Jornadas ITE - 2015 -Facultad de Ingeniería – UNLP
- [7] J. Osio, J. Salvatore, E. Kunysz, D. Montezanti, D. Alonso, V. Guarepi, M. Morales, "Análisis de Eficiencia en Arquitecturas Multiprocesador para Aplicaciones de Transmisión y Procesamiento de Datos", ITBA, CABA, WICC 2017
- [8] D. Montezanti, A. De Giusti, M. Naiouf, J. Villamayor, D. Rexachs, E. Luque, "A Methodology for Soft Errors Detection and Automatic Recovery", in Proceedings of the 15th International Conference on High Performance Computing & Simulation (HPCS). ISBN: 978-1-5386-3250-5/17. IEEE, 2017, pp. 434

- [9] D. Montezanti, E. Frati, D. Rexachs, E. Luque, M. Naiouf, and A. De Giusti, "SMCV: a methodology for detecting transient faults in multicore clusters," *CLEI Electronic Journal*, vol. 15, no. 3, pp. 5–5, 2012.
- [10] Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing". Pearson Addison Wesley, 2003.
- [11] F. Cappello, A. Geist, W. Gropp, S. Kale, B. Kramer, and M. Snir, "Toward exascale resilience: 2014 update," *Supercomputing frontiers and innovations*, vol. 1, no. 1, pp. 5–28, 2014.

“Tecnologías de la información y las comunicaciones mediante IoT para la solución de problemas en el medio socio productivo”

Jorge R. Osio^{1,2}, Juan. E Salvatore¹, Daniel Alonso¹, Valentin Guarepi¹, Marcelo Cappelletti¹, María Joselevich¹, Martin Morales¹

¹Instituto de Ingeniería y Agronomía - UNAJ

²UIDET CeTAD –Fac. de Ingeniería - UNLP

{josio, jsalvatore, dalonso, vguarepi, mcappelletti, mjoselevich, martin.morales}@unaj.edu.ar

Resumen

El presente proyecto se basa en la utilización de internet de las cosas (IoT) como herramienta fundamental para proveer soluciones tecnológicas a problemáticas de interés social como son la crisis energética, la ausencia de soluciones tecnológicas a sectores productivos marginados, el impacto del alto costo de los servicios básicos y el diseño de interfaces para educación a distancia. Entre las líneas de investigación que se llevarán adelante para contribuir con las problemáticas planteadas, se propone en primer lugar, atender las necesidades de sectores tales como el frutícola, hortícola y florícola, que carecen de soporte y no disponen de herramientas para automatizar procesos y acceder a información crítica en tiempo real mediante IoT. Por otro lado, se plantea el diseño de un algoritmo inteligente de eficiencia energética, que permita minimizar el consumo en sistemas alimentados con energía convencional y energía alternativa. Por último, mediante las interfaces IoT, se colaborará en el diseño de laboratorios remotos, propuestos en el proyecto de investigación dirigido por la Dra. María Joselevich, y en el control de un robot humanoide desarrollado específicamente

para acceder a lugares nocivos para la salud.

Palabras clave: *redes de sensores, Arquitecturas de procesadores embebidos, Internet de las cosas (IoT), Eficiencia Energética, Laboratorios Remotos, Robótica.*

Contexto

Las líneas de Investigación descriptas en este trabajo forman parte del Proyecto de Investigación Científico-Tecnológico “Tecnologías de la información y las comunicaciones mediante IoT para la solución de problemas en el medio socio productivo”, que se desarrolla en la Universidad Nacional Arturo Jauretche (UNAJ) y es continuación del proyecto presentado en la edición anterior de WICC [1].

El proyecto cuenta además, con financiamiento en el marco del programa “Universidad, Diseño y Desarrollo Productivo” del Ministerio de Educación a través del proyecto “Sistema de eficiencia energética” y está a la espera de su Aprobación para obtener financiamiento de la convocatoria “UNAJ Investiga 2017”.

Parte de las líneas de investigación desarrolladas se encuentran enmarcadas

en los convenios de colaboración en Actividades de Investigación firmados por la UNAJ con la UIDET-CeTAD y El convenio firmado con la Asociación de Productores Hostí-Florícolas de Varela y Berazategui APHOVABE.

1. Introducción

El concepto de internet de las cosas ha adquirido gran relevancia en los últimos años, debido a la posibilidad que ofrece de interconectar objetos entre si y la conectividad a internet que provee a las redes de objetos [10]. Aunque los sistemas IoT son más que conocidos, su implementación y desarrollo no ha sido completamente explotado en determinados sectores. IoT aplicado a la agricultura es un área en la que la evolución tecnológica nunca ha sido una prioridad, entre los factores que han postergado la investigación en esta área se tiene la falta de capacitación y falta de soporte e interacción con los productores para determinar cuáles son las principales problemáticas que se pueden resolver. Para este tipo de ambientes se dispone de un conjunto de protocolos de comunicación e interfaces como GSM/GPRS, Wifi, bluetooth y zigbee que se pueden implementar mediante herramientas basadas en Arquitecturas de Sistemas Embebidos como las desarrolladas en [5]. Las investigaciones realizadas mediante IoT se podrían aplicar en el monitoreo de cultivos, herramientas de soporte para la toma de decisiones, controlar automáticamente riego, protección de heladas, fertilización y monitoreo e iluminación artificial, entre otras.

Otra problemática común en el sector fruti-hortícola, es el acceso al suministro de energía, en cuyo caso requiere de sistemas costosos de energía alternativa.

Actualmente, se considera una problemática a tratar de manera urgente el uso eficiente de la energía [11], debido a que se está investigando mucho sobre energías alternativas pero hay mucho por hacer respecto a la reducción en el consumo de energía en función del factor humano y del control automático de los sistemas eléctricos/electrónicos. En este sentido, se debe tener en cuenta la necesidad de minimizar el consumo en los sistemas de energías alternativas, principalmente energía fotovoltaica [12], debido a que estos sistemas almacenan la energía en baterías altamente contaminantes y al lograr optimizar el consumo eficiente de energía se extendería notablemente la vida útil de la misma. Estos sistemas son de gran interés, principalmente por ser de uso común en los sectores agrícolas, donde se aplicarán las investigaciones realizadas conjuntamente con el proyecto “Estudio de sistemas inteligentes para optimizar el aprovechamiento de la radiación solar en la actividad agroindustrial del territorio de la UNAJ” dirigido por el Dr. Marcelo cappelletti. Para minimizar el consumo, se considera necesario implementar un algoritmo inteligente que tenga en cuenta todos los factores intervinientes, para esto es necesario diseñar un algoritmo basado en redes neuronales.

Respecto a la aplicación de IoT en robótica, tiene como objetivo el diseño y fabricación de un robot humanoide que posibilite el monitoreo remoto [13]. Los resultados de esta línea de investigación se orientarán a la realización de tareas en ambientes nocivos para el ser humano, como por ejemplo en la industria nuclear como medida de protección radiológica. Se prevé que el uso de sistemas robóticos se intensifique en los próximos años, es por eso que se considera necesario diseñar, desarrollar y construir nuevos sistemas que brinden soluciones a

situaciones problemáticas de manera eficiente y en la medida de lo posible, de bajo costo.

Respecto a la aplicación de IoT en Laboratorios remotos, desde el proyecto se diseñará la interfaz de Hardware y Software embebido entre servidor y los laboratorios desarrollados en el proyecto de la Dra. María Joselevich [14].

Metodología:

Para optimizar el uso racional de la energía, primero se deben tener en cuenta todos los factores de consumo, entre ellos el factor humano, el factor ambiental y el destino de cada uno de los recursos. Con un sistema diseñado a medida se puede lograr optimizar el consumo al máximo mediante la implementación de un algoritmo de decisión, basado en el perceptrón, que pueda interpretar los parámetros de entrada y generar una salida eficiente. Para tener en cuenta todos estos factores se debe desarrollar un sistema complejo que tenga la capacidad de combinar todos estos aspectos de manera óptima e inteligente.

En cuanto a las soluciones tecnológicas en las actividades florihortícolas en su mayoría requieren de la disponibilidad oportuna de datos, tanto del clima como del suelo, la iluminación y el agua. El desarrollo del sistema informático se realizará a partir de un algoritmo de software con soporte para diferentes aplicaciones, que en base a la información proporcionada por una red de sensores, emita mensajes de alertas a los usuarios y provea información de interés. Para que la propuesta sea aplicable en cualquier lugar, se plantea la utilización de una interfaz GSM/GPRS que permite transmitir información mediante las antenas celulares. El diseño del sistema

local a utilizar en huertas a campo abierto y bajo cubierta, estará formado por un sistema de procesamiento (basado en tecnologías de microcontroladores, arduino, LPC1769 y raspberry Phy), al cual se conectarán los sensores y el dispositivo de comunicación GSM/GPRS.

Para el diseño del robot humanoide, se propone utilizar sistemas embebidos de última generación, los cuales poseen alto poder de cómputo y bajo costo. Se implementarán sobre estos sistemas embebidos los algoritmos de automatismo y control necesarios para ejecución de las acciones del sistema robótico.

Antecedentes del Grupo de Trabajo

Se pretenden encausar temas de estudios de posgrado actuales y futuros de docentes y profesionales surgidos de la UNAJ dentro de estas líneas de investigación. El director del proyecto participa en proyectos de investigación desde el año 2005 en la Facultad de Ingeniería de la UNLP como colaborador y a partir de 2010 como investigador. Actualmente, el grupo de trabajo investiga en los temas: Trasmisión de datos IoT (mediante la interfaz Ethernet y GSM/GPRS), Eficiencia Energética, Robot Humanoide y Arquitecturas de procesadores embebidos; pertenecientes a estudios de Especializaciones, Maestrías y Doctorados que se realizan en el marco de acuerdos de colaboración entre el Laboratorio CeTAD (UNLP) y la UNAJ a través del proyecto de investigación.

Como resultado del trabajo realizado recientemente en la UNAJ se han publicado artículos en diferentes congresos nacionales de Ingeniería Informática y Ciencias de la Computación [1 - 5].

2. Líneas de Investigación y Desarrollo

El grupo de investigación que se ha constituido en la UNAJ es multidisciplinario, y sus miembros cuentan con experiencia en sistemas multiprocesador, Arquitecturas de Procesadores Embebidos, IoT, sistemas operativos embebidos, protocolos de comunicación TCP y UDP y robótica.

En cuanto a la eficiencia energética, en la actualidad se considera que un sistema de iluminación es eficiente energéticamente cuando permite adaptar el nivel de iluminación en función de la variación de la luz solar o en función de la detección o no de presencia. En esta propuesta se considera que las soluciones existentes no ha sido suficientemente desarrollada, ya que para un uso eficiente de la energía se deben tener en cuenta varios factores y combinar una amplia variedad de métodos para lograr un algoritmo inteligente que permita optimizar el consumo al máximo.

En cuanto a la investigación aplicada al sector fruti-hortícola, se plantea el diseño de una red de sensores que interactúan entre si y se comunican con un servidor mediante una red IoT, en este sentido se propone usar el protocolo LoRa para determinar la eficiencia de dicho protocolo en aplicaciones que exigen el uso energía alternativa.

En cuanto a la robótica, el robot humanoide ha tenido poco desarrollo en Latinoamérica y la investigación orientada a inspeccionar sectores de difícil acceso o en ambientes donde la contaminación impide la permanencia de seres humanos en un nicho poco explotado. Para esto se deberá desarrollar algoritmos de aprendizaje basados en redes neuronales básicas y monitoreo remoto mediante IoT.

Temas de Estudio e Investigación

- Programación de diferentes Arquitecturas de Procesadores (freescale, LPC1769, CIAA, Arduino, Raspberry Phy) embebidos para la implementación de nodos en las redes de sensores y para la implementación de algoritmos de control.
- Utilización y configuración de sistemas operativos de tiempo real y Linux embebidos para la ejecución de los algoritmos y la adquisición de datos de sensores.
- Implementación de una red de sensores inalámbrica mediante protocolos estándar de IoT (LoRA) e interfaces RF.
- Investigación relacionada con los protocolos TCP y UDP ([6] y [7]) para el envío de paquetes, donde, en condiciones de red conocidas, se deben determinar las tasas de transmisión mínimas para asegurar el funcionamiento óptimo del sistema de eficiencia energética y los laboratorios remotos ([8] y [9]).
- Determinación de la eficiencia en la ejecución de algoritmos de control de un robot humanoide sobre un sistema multicore.
- Implementación de un algoritmo de básico de redes neuronales para la optimización del consumo en los sistemas de eficiencia energética.

3. Resultados Obtenidos/Esperados

Hasta el momento se han obtenido resultados satisfactorios en relación a los siguientes temas:

- Se obtuvieron resultados aceptables en la implementación de una

plataforma para IoT basada en el procesador Cortex M3 de 32 bits y la interfaz wifi.

- Se implementó un sistema de telemetría mediante la interfaz GSM/GPRS mediante el módulo M92 de Quectel, que permite supervisar el estado del cultivo (temperatura, iluminación y humedad) mediante una aplicación móvil.
- Por último, se encuentra en la etapa final un sistema que permite optimizar el consumo energético en instituciones educativas mediante la ejecución de algoritmos que toman en cuenta diferentes factores. Este sistema utiliza la plataforma IoT desarrollada en [5] para la implementación y ejecución del algoritmo.

Los resultados esperados para los nuevos temas que se están desarrollando en el proyecto son:

- Se espera para el año en curso realizar las pruebas de funcionamiento en campo del sistema autónomo y escalable que se encarga de minimizar el consumo de energía conjuntamente con la implementación de un sistema informático que provea información precisa de magnitudes climatológicas, del estado del suelo y del nivel de iluminación para determinados cultivos. En función de los resultados que se obtengan se realizarán las correcciones pertinentes.
- El objetivo en cuanto al sistema robótico, consiste en finalizar el diseño y programación del robot humanoide que provea soluciones innovadoras, eficientes y de bajo

costo para inspección en ambientes inhóspitos.

- Finalmente, la incorporación de esta nueva línea buscará proveer soluciones de software embebido para la implementación de laboratorios remotos propuestos en el Proyecto de la Dra. María Joselevich.

4. Formación de Recursos Humanos

Dentro de la temática de la línea de I+D, todos los miembros del proyecto participan en el dictado de asignaturas de la carrera de Ingeniería Informática de la UNAJ.

En este proyecto existe cooperación a nivel nacional. Hay un investigador realizando Doctorado, tres realizando Maestrías en temas relacionados con sistemas multiprocesador, IoT, sistemas embebidos y software embebido y transmisión de datos mediante TCP y UDP. Además, participan tres Doctores especializados en la temática y se cuenta con la colaboración de estudiantes avanzados.

5. Bibliografía

- [1] J. Osio, J. Salvatore, E. Kunysz, D. Montezanti, D. Alonso, V. Guarepi, M. Morales, "Análisis de Eficiencia en Arquitecturas Multiprocesador para Aplicaciones de Transmisión y Procesamiento de Datos", ITBA, CABA, WICC 2017.
- [2] J. Osio, D. Montezanti, M. Morales, "Análisis de Eficiencia en Sistemas Paralelos", Ushuahia, Tierra del Fuego, WICC 2014
- [3] J. Osio, J. Salvatore, M. Morales, "Análisis de Eficiencia en Arquitecturas Multiprocesador en aplicaciones de

- Telemedicina”, UNS, Ciudad de Salta, WICC 2015.
- [4] J. Osio, J. Salvatore, E. Kunysz, V. Guarepi, M. Morales, “Análisis de Eficiencia en Arquitecturas Multiprocesador para Aplicaciones de Transmisión y Procesamiento de Datos”, UNER, Ciudad de Concordia, WICC 2016.
- [5] Osio J., C. Aquarone, E. Hromek, J. Salvatore, “Plataforma de desarrollo para IoT”, IV conaiisi, 2017
- [6] Richard Stevens, “The Protocols TCP/IP Illustrated, Volume 1”, 1993
- [7] Jan Axelson, “Embedded Ethernet a internet complete”, LLC, 2003
- [8] L. Iogna Pratt, J. Osio, J. Rapallini, “Implementación de un web server embebido”, Fac. Ing., UNLP, 2013.
- [9] J. Salvatore, “Desarrollo de un Sistema de voz sobre IP”, Fac. Ing., UNLP, 2012.
- [10] Leila Fatmasari Rahman, ”Choosing your IoT Programming Framework: Architectural Aspects”, 2016 IEEE 4th International Conference on Future Internet of Things and Cloud
- [11] Eficiencia energética en los edificios, José María Fernández Salgado, AMV ediciones, 2011
- [12] Energía Solar Fotovoltaica 2ª Edición, Miguel Aparicio, Edit. marcombo, 2009
- [13] Mohammad Jamshidi y Patrick J. Eicker, Robotics and remote system for hazardous environments, Prentice Hall, New Jersey, Estados Unidos (1993)
- [14] Laboratorio remoto en un entorno virtual de enseñanza aprendizaje, M. A. Revuelta, FI- UNLP, 2016

Contienda entre las variantes del protocolo TCP Vegas y Reno por los recursos de la red en un modelo híbrido simple

Diego R. Rodríguez Herlein, Carlos A. Talay, Claudia N. González y Franco

A. Trinidad

Campus Universitario – Oficina 18/ Dpto. Ciencias Exactas e Informática UARG / UNPA
{dherlein, ctalay, cgonzalez}@uarg.unpa.edu.ar, tfrancoalejandro@gmail.com

Luis A. Marrone

L.I.N.T.I. – Universidad Nacional de La Plata
Calle 50 y 120 – 2 do. Piso – Edificio Bosque Oeste
lmarrone@linti.unlp.edu.ar

Resumen

El presente trabajo aborda la problemática que se observa al convivir diferentes variantes de los algoritmos de control de congestión en la competencia por los recursos compartidos de la red. Dada la naturaleza de las redes IP, es deseable el uso justo de estos recursos por los distintos flujos TCP, sin perder de vista las posibles degradaciones del rendimiento en los enlaces inalámbricos. De esta manera, se propone analizar cómo interactúan dos flujos TCP similares que utilizan algoritmos de control de congestión conceptualmente diferentes, como son Reno y Vegas, compitiendo por los recursos de una red híbrida con enlaces cableados e inalámbricos

Palabras clave: TCP, red híbrida, contienda, ancho de banda

Contexto

Este estudio está enmarcado en el proyecto de investigación 29/A396 “Evaluación de desempeño del protocolo TCP en topologías mixtas cableadas-inalámbricas” radicado de la UNPA-

UARG, continuación del proyecto 29/A358-1 “Análisis de performance del protocolo TCP utilizado en redes móviles”. El proyecto está compuesto mayoritariamente de docentes de la UNPA-UARG, dirigido por el Sr. Carlos A. Talay y cuanta como Co-director al Sr. Luis A. Marrone perteneciente a la UNLP y se financia íntegramente con fondos destinados a proyecto de investigación de la UNPA-UARG.

Introducción

El protocolo TCP [1] (Transmission Control Protocol) es el encargado de transportar la mayor parte del tráfico de internet, es por ello que el rendimiento de internet depende, en gran medida, de cómo funciona TCP. Las características de rendimiento de una versión particular de TCP se definen por el algoritmo de control de congestión que implementa. En las redes que incluyen enlaces inalámbricos, el control de congestión puede ser aún más determinante en el rendimiento de este protocolo. En el presente proyecto de investigación se analizan algunas de las propuestas de

control de congestión para mejorar el rendimiento pero que preservan su principio fundamental de comunicación host a host, es decir, que no dependen de ningún tipo de señalización explícita de la red.

La característica fundamental de TCP es su capacidad para proporcionar un canal fiable, bidireccional y virtual entre dos hosts en Internet. Dado que el protocolo funciona a través de la red IP, es TCP el que debe proporcionar esta fiabilidad, implementando un control de flujo basado en ventana deslizante. [2]

Desafortunadamente, el TCP original carecía de algún medio para ajustar la velocidad de transmisión de acuerdo al estado de la red. Esto crea varios efectos inesperados, siendo el más importante el colapso por congestión. Por ello el uso inteligente de los recursos disponibles en las redes de conmutación de paquetes, no es un problema trivial.

Para resolver el problema del colapso de congestión, se han propuesto varias soluciones, donde todas ellas comparten la misma idea: la introducción de un mecanismo que limite para la tasa de envío. Para ello se introdujo el concepto de ventana de congestión, cuyo propósito es estimar de la cantidad de datos que la red puede aceptar para su entrega sin que se produzca congestión.

Internet es un entorno heterogéneo y el uso efectivo de la red no sólo depende de que un único flujo TCP pueda utilizar la capacidad de la red, sino también de qué tan bien coopera con otros flujos a través

de la misma red. La eficacia no es el único parámetro importante de los algoritmos de control de congestión, también deben hacer valer el uso justo de recursos compartidos.

Es por ello que el análisis de dos flujos de datos, regidos por el protocolo TCP, pero con distintas variantes, puede aportar una visión acerca de cómo es utilizado el ancho de banda bajo distintas estrategias, y ver cómo interactúan entre sí ante una competencia por la utilización de este recurso. Explorando este camino, planteamos aquí un modelo simple, con dos flujos que compiten entre sí en la misma trayectoria de red en una topología híbrida, sin otros flujos presentes.

El caso de estudio

Los algoritmos de control de congestión, tratan de determinar dinámicamente el ancho de banda disponible y la latencia de la red. Luego en función de su análisis se modificar la tasa de envío del emisor TCP para evitar el colapso de la subred.

El control de congestión [3] comienza con el algoritmo Slow Start, en donde se incrementa la ventana de congestión (cwnd) en forma exponencial [4] hasta que se alcanza el valor del umbral (ssthresh). A partir de este punto, comienza el algoritmo de Congestion Avoidance [5], que incrementa el valor de cwnd en forma lineal. Si se retrasa el ACK de un paquete o bien recibe un ACK duplicado, TCP asume que el paquete se perdió y lo retransmite. Es en este momento que se activa el algoritmo

de retransmisión rápida (Fast Retransmit) [6].

Dentro de las implementaciones de TCP existentes [7] se analizaran:

TCP Reno [8]: Cuando se detecta una pérdida de paquete, TCP Reno dispara el algoritmo de Fast Recovery [9], reenviando el paquete perdido y reduciendo el umbral ssthresh a la mitad, evitando disparar el algoritmo Slow Start. Cuando Reno sale del algoritmo Fast Recovery, da paso al algoritmo Congestión Avoidance, lo que permite que se recupere más rápido de la congestión.

TCP Vegas [10], [11]: Hasta ahora la ventana de congestión crecía hasta que ocurría una pérdida de paquete. TCP Reno asume que hay congestión cuando se detectan estas pérdidas. La idea del TCP Vegas es controlar y mantener el tamaño adecuado de la ventana de manera de no llegar a la pérdida de paquetes. El TCP Vegas administra el tamaño de la ventana observando el RTT (round-trip-time) de los paquetes que el emisor envió con anterioridad. Si el RTT crece, TCP Vegas reconoce que la red comienza a estar congestionada y reduce su ventana de congestión. Por el contrario, si los valores de RTT decrecen, el emisor determina que la red no está congestionada e incrementa el tamaño. Para evitar las oscilaciones, se define un intervalo en donde se mantiene el tamaño de la ventana.

TCP fue desarrollado y posteriormente optimizado para redes cableadas, lo que

implica que la pérdida de paquetes se debe casi con exclusividad a la congestión en la red, suponiendo una tasa de error de bits en tránsito prácticamente inexistente. Con el crecimiento en la demanda de movilidad, se comenzó a utilizar en enlaces inalámbricos. Estos enlaces tienen características muy diferentes, de manera que los retrasos y las pérdidas de paquetes, ya no se deben exclusivamente a la congestión, sino también al daño de paquetes en tránsito [12] y la desconexión temporal de los nodos, lo que puede ocasionar un comportamiento muy poco eficiente del protocolo en ese entorno. Esta complejidad adicional se contempla en el modelo del presente artículo, al utilizar una topología híbrida (cableado e/inalámbrico), lo que plantea un caso más general de análisis.

Diseño del modelo de prueba

El modelo implementado en el simulador NS-2 (Network Simulator 2, v2.35) representa una topología híbrida de 5 nodos (2 nodos cableados, 1 estación base y 2 nodos inalámbricos), como se observa en la siguiente figura (Fig.1)

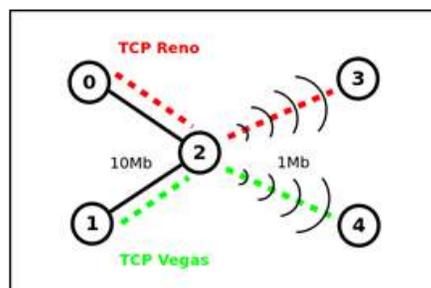


Figura 1. Modelo de estudio

Los nodos 0 y 1 están conectados a la estación base (nodo 2) mediante enlaces cableados full dúplex, 10Mb/s, retardo de propagación de 2ms y política de cola DropTail. Para el enlace inalámbrico se establece un ancho de banda de 1 Mb/s, modo de propagación TwoRayGround, capa física WirelessPhy, MAC 802.11, antena OmniAntena y los dos nodos (3 y 4) no poseen movimiento.

Se realizaron tres pruebas con dos tráficos FTP de 3.000 paquetes de 1.000 bytes cada uno, utilizando las variantes de TCP Reno y Vegas. En todos los casos, el flujo del TCP Reno es desde el Nodo 0 al Nodo 3 y para el TCP Vegas del Nodo 1 al Nodo 4,

En la primera simulación, ambos flujos TCP comienzan simultáneamente, a los 5 seg. de iniciado el ensayo (fig. 2). En la segunda corrida, solo el flujo de TCP Vegas comienza a los 5 seg. y se demora hasta los 10 seg. el comienzo del flujo TCP Reno (fig. 3). En la tercera simulación se invierten los flujos, y es TCP Reno quien comienza a transmitir a los 5 seg. y TCP Vegas a los 10 seg. (fig. 4).

De cada una de las prueba se obtuvo throughput instantáneo (cantidad de Bytes en tránsito en la red en determinado momento), utilizando un script AWK que filtró los valores correspondientes del archivo de traza generado por NS2. Luego se utilizó gnuplot para graficar los resultados.

Resultados obtenidos y futuras líneas de investigación

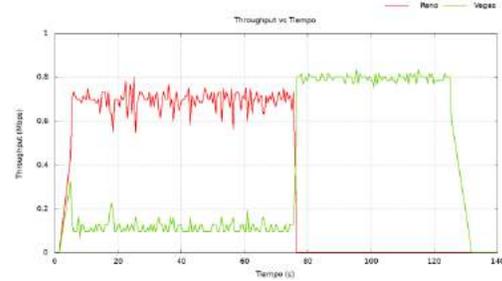


Figura 2. Reno y Vegas inicio simultáneo

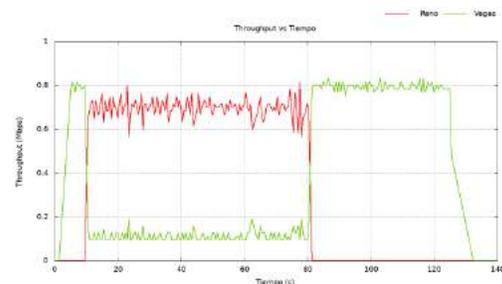


Figura 3. Inicia Vegas y luego Reno con 5 seg. de retraso

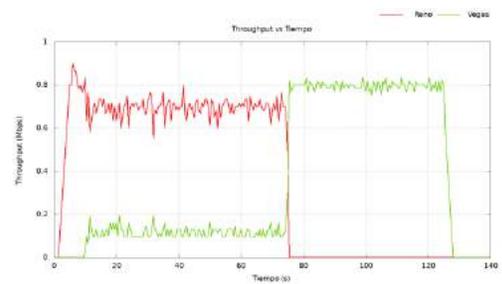


Figura 4. Inicia Reno y luego Vegas con 5 seg. de retraso

En la confrontación por recursos en un esquema de transmisión bajo estas condiciones son muy claras. La agresividad del protocolo TCP Reno acapara gran parte del ancho de banda al momento de competir contra su rival, en este caso el protocolo TCP Vegas. En la Figura 2, vemos un ensayo donde las transmisiones comienzan en forma simultánea y desde el comienzo Reno captura el mayor ancho de banda posible, relegando al protocolo Vegas al desarrollo de un tráfico mínimo, retrasando notablemente su ritmo de

transmisión de datos. Esto se confirma en cuanto, comenzando al mismo tiempo y habiendo transmitido el mismo volumen de datos, para el caso de TCP Reno emplea 77 seg. para completar el envío, cuando para el protocolo Vegas se tiene 132 seg. De la misma forma, en las Figuras 3 y 4, cuando los comienzos de los flujos no son simultáneos, se puede observar idéntico comportamiento.

En base al ejemplo planteado, vemos que si por competencia de recursos de tratase, la variante más agresiva es la que domina la competencia en la transmisión de datos. La agresividad en el mecanismo de transferencia de datos marca la diferencia y es el factor distintivo que hace que un protocolo domine por sobre otro. Esto se verifica además si analizamos el tiempo total de transmisión de los datos. A pesar de ello hay un hecho a resaltar, si consideramos el nivel instantáneo de throughput, observamos que cuando el protocolo Vegas se encuentra transmitiendo solo en el enlace, su valor promedio es más alto, por tanto podemos inferir que posee un mejor aprovechamiento de los recursos disponibles de la red.

Si bien la prueba es concluyente, ampliarse el estudio con ensayos donde se presente otras topologías con un mayor espectro de variantes del protocolo TCP y estudiar si se puede introducir propuestas que brinden más equidad al funcionamiento de la red.

Formación de Recursos Humanos

El grupo de investigación conformado se caracteriza una constitución heterogénea de profesionales vinculados a la informática. Entre ellos podemos enumerar una Licenciada en Informática, un Ingeniero en Electrónica y un Ingeniero Electricista con un máster en Sistemas y Redes de Comunicaciones. También integra el grupo dos alumnos de la carrera Lic. en Sistemas de la UNPA-UARG. En el transcurso del proyecto se tiene como objetivo consolidar la formación en investigación de los integrantes de menos antecedentes en proyectos y también está contemplado que uno de los integrantes complete su trabajo final de maestría.

Referencias

- [1] Postel J., "RFC 793: Transmission Control Protocol", September 1981.
- [2] M. Handley, J. Padhye and S. Floyd, "TCP Congestion Window Validation", RFC 2861, June 2000.
- [3] V. Jacobson, "Congestion Avoidance and Control," ACM SIGCOMM '88, Vol. 118 N°4, August 1988.
- [4] Stevens, W., "TCP slow start, congestion avoidance, fast retransmit, and fast re-recovery algorithms". RFC 2001, 1997.
- [5] Jacobson V., "Congestion Avoidance and Control", ACM SIGCOMM Computer Communication Review, 25(1), 1995, pp. 157-187.
- [6] W. Stevens, "RFC2001—TCP Slow Start, Congestion Avoidance, Fast Retransmit," RFC, 1997
- [7] Saleem-ullah Lar, Xiaofeng Liao, "An initiative for a classified bibliography on TCP/IP congestion control", Journal of Network and Computer Applications 36, 126–133, 2013
- [8] V. Jacobson, "Modified TCP congestion avoidance algorithm," email to the end2end list, April 1990.
- [9] M. Allman, V. Paxson, and W. Stevens, "RFC2581—TCP congestion control" RFC, 1999.
- [10] L. Brakmo and L. Peterson, "TCP Vegas: end to end congestion avoidance on a global Internet," IEEE Journal Selected Areas Communications. vol. 13, no. 8, pp. 1465–1480, October 1995.
- [11] G. Hasegawa, K. Kurata, and M. Murata, "Analysis and improvement of fairness between TCP Reno and Vegas for deployment of TCP Vegas to the Internet" in Proc. IEEE ICNP, 2000, pp. 177–186.
- [12] Elaarag, H., "Improving TCP Performance over Mobile Networks", ACM Computing Surveys, Vol. 34, No. 3, September 2002.

Evaluación de performance en Redes Definidas por Software

Diego Bolatti, Ricardo Calcagno,
Carlos Cuevas, Sergio Gramajo, Reinaldo Scappini

Grupo Ingeniería en Sistemas de Información, Universidad Tecnológica Nacional Facultad
Regional Resistencia,
French 414, (3500) Chaco, Argentina. Tel. 362-4432683
{dbolatti, rcalcagno, cac, sergio, [rscappini](mailto:rscappini@frre.utn.edu.ar)} @frre.utn.edu.ar

Resumen

El objetivo principal de este artículo es presentar una metodología sencilla para la simulación de una SDN (Software Defined Networking) [1], configurada con el protocolo OpenFlow [2] controladores POX [3] y ODL [4], utilizando la herramienta de simulación Mininet [5], y realizar pruebas de performance mediante la utilización de IPERFv3 [6], mostrando los resultados con el software GNUPlot [7], todos ellos con licencia GNU. Además, se describen los requerimientos y recomendaciones fundamentales para dicha emulación, y se presentan los resultados de pruebas sencillas con el fin de verificar la conectividad, transferencia de datos y operación sobre una topología de prueba. Se orienta al desarrollo de modelos de evaluación de performance que ayuden a los administradores de red a tomar decisiones en base a atributos críticos (tipos de tráfico, servicios, usuarios, etc.) identificados previamente.

Palabras clave: Redes Definidas por Software, Análisis de Tráfico, Performance

Contexto

Este proyecto está inserto en una línea de I/D presentada en la Universidad Tecnológica Nacional con código: UTN-2422. Título: “Modelo para la evaluación de performance mediante identificación de tráfico y atributos críticos en Redes Definidas por Software”. Dicho proyecto se lleva a cabo en el ámbito del Dpto. de Ingeniería en Sistemas de Información perteneciente a la Facultad

Regional Resistencia de la Universidad Tecnológica Nacional.

Introducción

Si bien el objetivo del presente trabajo no radica en la descripción de tecnologías aplicadas en SDN, se brinda un breve contexto descriptivo del ámbito de trabajo, a los efectos de una mejor comprensión. Las Redes Definidas por Software proponen un modelo para cubrir nuevas demandas de usuarios y organizaciones. Ésta es una arquitectura de red emergente, donde el control de la infraestructura de red está desacoplado del reenvío de datos y, a su vez, es directamente programable. La inteligencia de red es (lógicamente) centralizada en controladores SDN basados en software que mantienen una visión global de la red. Como resultado, las organizaciones controlan la red independiente del proveedor en un único punto lógico lo que simplifica, en gran medida, el diseño de la red y su operación. A su vez, se simplifica la gestión de los dispositivos de red debido a que ahora no se tienen que entender y procesar múltiples normas de protocolo, sino simplemente aceptar instrucciones de los controladores centralizados de SDN (Fig. 1).

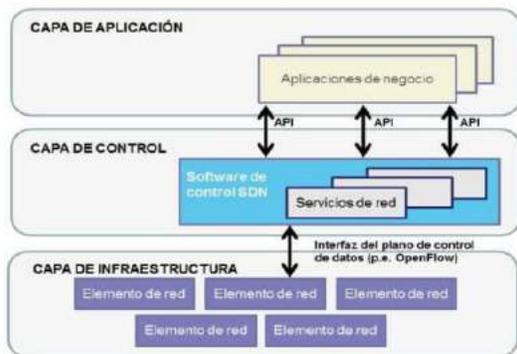


Figura 1. Arquitectura de Software Defined Network

Tal vez lo más importante de esta nueva visión de red, es que se pueda configurar mediante programación esta abstracción simplificada de la red en lugar de tener que configurar manualmente múltiples dispositivos. Además, aprovechando la inteligencia centralizada del controlador SDN, se puede alterar el comportamiento de la red en tiempo real y desarrollar nuevas aplicaciones y servicios de red ágilmente, lo que mejora sustancialmente las posibilidades, a esto se lo conoce por Northbound Interfaces [8]: Una interfaz hacia el norte o capa de aplicación, la cual sigue siendo una cuestión abierta, es muy pronto para estandarizarla, se basará en la experiencia con los diferentes controladores. Además, se debe permitir que las aplicaciones de red no dependan de implementaciones específicas. Para poder concretar la nueva arquitectura de redes fue necesario crear y estandarizar una interfaz de comunicaciones entre el control y el reenvío de datos, conocida por Southbound Interfaces [8]: Son los puentes de conexión entre los elementos de control y los de reenvío, siendo el elemento esencial para separar la funcionalidad del plano de datos y de control. Las API están todavía muy enlazadas a dispositivos físicos de reenvío de paquetes en una infraestructura física. Un switch puede tardar bastante tiempo de desarrollo para estar presente en el mercado, OpenFlow es uno de los primeros estándares de interface de comunicación definida entre las capas de control y forwarding en una arquitectura SDN. Además, permite el acceso directo y la manipulación de la capa de forwarding de los dispositivos de Red (switches, routers), ya sean físicos o virtuales (basadas en un Hypervisor)..

Para ello se creó el protocolo OpenFlow que permite el acceso directo a la gestión de datos de reenvío en dispositivos de red como switches y routers, tanto físicos como virtuales y de un modo abierto. Esto contrasta con las arquitecturas de redes tradicionales donde los dispositivos de red son monolíticos y cerrados. Ningún otro protocolo estándar tiene la funcionalidad y finalidad de OpenFlow que transfiere el control de los dispositivos de red a la lógica del software de control.

Una característica particular de SDN basada en OpenFlow es que se puede implementar en las redes existentes, tanto físicas como virtuales. Los dispositivos de red pueden realizar el reenvío basado en OpenFlow, así como el reenvío tradicional, lo que hace que sea muy fácil para las organizaciones introducir progresivamente esta tecnología, incluso en los entornos de red de múltiples proveedores.

Con este panorama es atendible la necesidad de contar con herramientas que permitan evaluar a priori el comportamiento de los cambios introducidos en las aplicaciones y servicios de las redes como también el impacto en la performance de éstas

Líneas de investigación y desarrollo

En el proyecto “*Modelo para la evaluación de performance mediante identificación de tráfico y atributos críticos en Redes Definidas por Software*” se propone el análisis de la arquitectura y estándar de SDN. Además del diseño e implementación de un sistema de soporte a las decisiones con Información Lingüística [9],[10] y [11] para evaluar datos **cualitativos** de SDN y su uso. Este sistema permite que múltiples expertos puedan participar conjuntamente dando sus puntos de vista que ayuden a la toma de decisiones.

Los aspectos **cuantitativos** del modelo se realizan mediante simulaciones y creación de escenarios de comparación con las redes tradicionales.

Referencias [1],[2],[3],[4],[5],[6],[7] y [8]

Resultados y objetivos

En este trabajo se han obtenido resultados utilizando los recursos mencionados arriba,

como se verá a continuación. Se utiliza Mininet: ágil y potente simulador de SDN para la creación de escenarios de prueba, teniendo en cuenta particularmente la posibilidad que ofrece en la caracterización de los parámetros de enlaces, tanto en lo que hace a las características del enlace (ancho de banda, retardo etc.) como a parámetros relacionados con calidad de servicio (Ver figura 2). Esto se puede realizar en forma sencilla creando las topologías con la aplicación Miniedit [12], incluida en Mininet, (ver figura 3), que permite tomar y arrastrar los elementos y definir sus propiedades con solo la ayuda del mouse. Lo bueno de Miniedit es que permite construir muy velozmente la topología en un entorno gráfico y genera automáticamente las scripts para utilizarlas con Mininet (Cabe mencionar que tanto en estas scripts como así también en la inmensa mayoría de las aplicaciones en estos entornos se utiliza Python [13]).

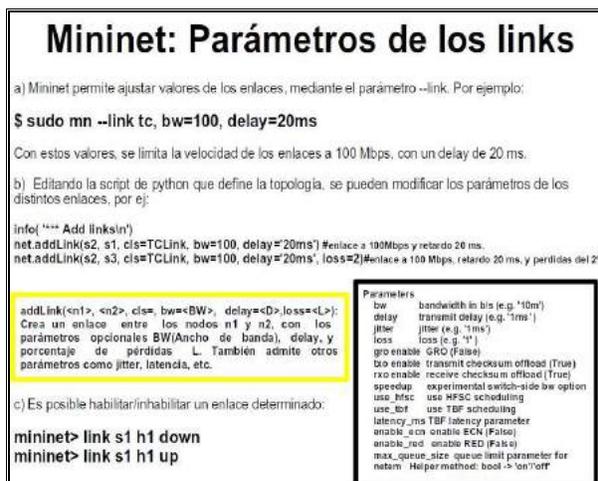


Figura 2 Ejemplos de configuración de parámetros de enlaces (propio).

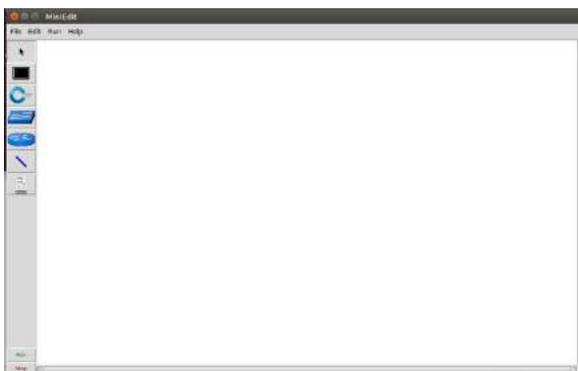


Figura 3: Interfaz gráfica de Miniedit (propio)

Ejemplo de topología de prueba: a los efectos de realizar una simulación y evaluar la performance se muestra una topología de prueba (figura 4). En la misma se muestran tres locaciones con redes lan en cada una y de las siguientes características:

- Enlaces LAN en las tres locaciones a 100 Mbps con un delay de 5 ms y porcentaje de pérdida de paquetes de 1%
- Enlace SW1 SW2 1000 Mbps con delay de 15 ms., porcentaje de pérdida de paquetes 1% y Jitter de 1%
- Enlace SW2 SW3 1000 Mbps con delay de 20 ms., pérdida de paquetes 1% y Jitter de 1%

Con respecto del controlador, se puede definir para la misma topología, un controlador POX, ODL, o bien cualquiera disponible solo basta con definir las propiedades haciendo un clic con el botón derecho sobre el ícono Controlador.

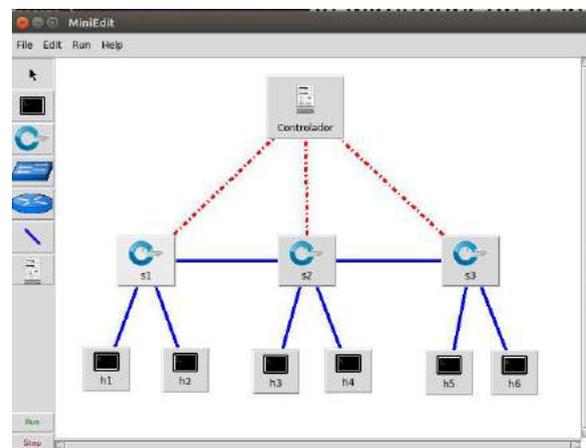


Figura 4: Topología de prueba (propio)

A partir de la topología definida; y ajustada la script con todos los detalles que deseemos, es posible iniciar las pruebas de performance utilizando IPERF3, para ello basta con abrir terminales en los hosts correspondientes y lanzar instancias ya sea de cliente o servidor según el/los enlaces que deseemos probar. Vale aclarar que IPERF3 tiene muchas mejoras y ventajas respecto a la versión original IPERF, las más destacables son la caracterización del tráfico tanto en TCP como

UDP, la posibilidad de lanzar múltiples instancias respecto de un mismo servidor, como así también crear streams en paralelo para cada instancia, se puede ver con detalle en [6]

A continuación, se muestra un ejemplo suponiendo un tráfico TCP de 80 Mbps entre h1 y h3 de la topología de la figura 4.

En h3 un servidor: `iperf3 -s -p 5566 -i 1`(fig.5)

```

"Node: h3"
root@mininet-vm:~# iperf3 -s -p 5566 -i 1
warning: this system does not seem to support IPv6 - trying IPv4
server listening on 5566
-----
Accepted connection from 10.0.0.1, port 56412
[ 27] local 10.0.0.3 port 5566 connected to 10.0.0.1 port 56414
[ ID] Interval      Transfer      Bitrate
[ 27] 0.00-1.00    sec  9.42 MBytes  79.0 Mbits/sec
[ 27] 1.00-2.00    sec  9.85 MBytes  80.2 Mbits/sec
[ 27] 2.00-3.00    sec  8.67 MBytes  72.7 Mbits/sec
[ 27] 3.00-4.00    sec  5.11 MBytes  42.3 Mbits/sec
[ 27] 4.00-5.00    sec  9.25 MBytes  77.6 Mbits/sec
[ 27] 5.00-6.00    sec  9.46 MBytes  79.4 Mbits/sec
[ 27] 6.00-7.00    sec  9.85 MBytes  82.6 Mbits/sec
[ 27] 7.00-8.00    sec  3.97 MBytes  33.3 Mbits/sec
[ 27] 8.00-9.00    sec  10.2 MBytes  85.9 Mbits/sec
[ 27] 9.00-10.00   sec  8.34 MBytes  70.0 Mbits/sec
[ 27] 10.00-11.00  sec  10.4 MBytes  87.3 Mbits/sec
[ 27] 11.00-12.00  sec  6.56 MBytes  55.1 Mbits/sec
[ 27] 12.00-13.00  sec  3.79 MBytes  31.8 Mbits/sec
[ 27] 13.00-14.00  sec  10.4 MBytes  86.9 Mbits/sec
[ 27] 14.00-15.00  sec  11.1 MBytes  93.5 Mbits/sec
[ 27] 15.00-15.03  sec  156 KBytes  37.0 Mbits/sec
-----
[ ID] Interval      Transfer      Bitrate
[ 27] 0.00-15.03   sec  126 MBytes  70.5 Mbits/sec
iperf3: the client has unexpectedly closed the connection
server listening on 5566

```

Figura 5: Servidor iperf3 (propio)

En h1 un cliente: `iperf3 -c 10.0.0.3 -t 15 -b 80M` (fig.6)

```

"Node: h1"
root@mininet-vm:~# iperf3 -c 10.0.0.3 -p 5566 -t 15 -b 80M
Connecting to host 10.0.0.3, port 5566
[ 27] local 10.0.0.1 port 56414 connected to 10.0.0.3 port 5566
[ ID] Interval      Transfer      Bitrate      Retr  Cwnd
[ 27] 0.00-1.00    sec  9.54 MBytes  80.0 Mbits/sec  133  52.3 KBytes
[ 27] 1.00-2.00    sec  9.62 MBytes  80.7 Mbits/sec  144  36.8 KBytes
[ 27] 2.00-3.00    sec  8.88 MBytes  74.3 Mbits/sec  149  24.0 KBytes
[ 27] 3.00-4.00    sec  5.12 MBytes  43.1 Mbits/sec  74  11.3 KBytes
[ 27] 4.00-5.00    sec  9.25 MBytes  77.5 Mbits/sec  90  74.9 KBytes
[ 27] 5.00-6.00    sec  9.50 MBytes  79.8 Mbits/sec  129  48.1 KBytes
[ 27] 6.00-7.00    sec  9.88 MBytes  82.5 Mbits/sec  168  46.7 KBytes
[ 27] 7.00-8.00    sec  4.00 MBytes  33.7 Mbits/sec  98  35.4 KBytes
[ 27] 8.00-9.00    sec  10.2 MBytes  86.0 Mbits/sec  137  42.4 KBytes
[ 27] 9.00-10.00   sec  8.38 MBytes  70.3 Mbits/sec  93  49.5 KBytes
[ 27] 10.00-11.00  sec  10.4 MBytes  87.0 Mbits/sec  180  35.4 KBytes
[ 27] 11.00-12.00  sec  6.50 MBytes  54.5 Mbits/sec  145  18.4 KBytes
[ 27] 12.00-13.00  sec  3.75 MBytes  31.5 Mbits/sec  64  19.8 KBytes
[ 27] 13.00-14.00  sec  10.5 MBytes  88.1 Mbits/sec  148  45.2 KBytes
[ 27] 14.00-15.00  sec  11.1 MBytes  93.3 Mbits/sec  107  74.9 KBytes
-----
[ ID] Interval      Transfer      Bitrate      Retr  sender receiver
[ 27] 0.00-15.00   sec  127 MBytes  70.8 Mbits/sec  1859
[ 27] 0.00-15.03   sec  126 MBytes  70.5 Mbits/sec
iperf Done.
root@mininet-vm:~#

```

Figura 6: Cliente iperf3 (propio)

Para documentar en forma de gráfico los resultados es útil la aplicación GNUPLOT que toma directamente los resultados de la salida de iperf y los convierte en una gráfica a la que se le pueden agregar múltiples detalles, ver referencias [14], [15]. El gráfico correspondiente al ejemplo se muestra en la figura 7.

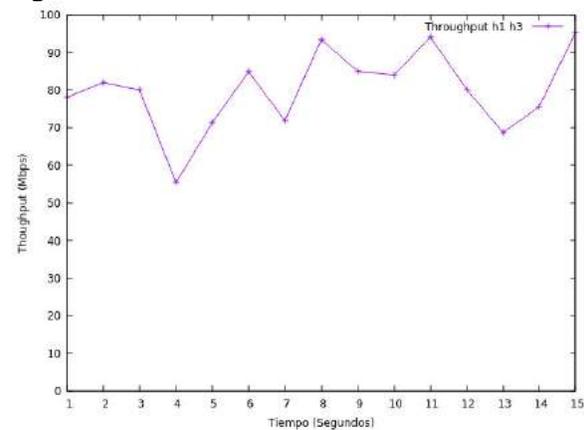


Figura 7: Grafico de throughput entre h1 y h3 en Mbp

Con la misma secuencia de operaciones, se pueden graficar múltiples instancias de simulación como así también cambiar muy fácilmente los parámetros de simulación, sencillamente cambiándolos en línea de comandos para luego visualizar los cambios con poco esfuerzo.

Formación de Recursos Humanos

La formación de recursos humanos es la siguiente:

- Formación de becarios:** Dos becarios alumnos avanzados de la carrera de ingeniería en sistemas de información y un becario graduado de iniciación a la investigación. Esto hará posible fomentar la actividad de investigación en alumnos que están próximos a recibirse y graduados jóvenes estimulando la actividad de investigación.
- Formación de postgrado:** A partir de las líneas de investigación desarrolladas en el proyecto se prevé que el Ing. Carlos Cuevas finalice su Maestría en Redes de la Universidad de La Plata

mediante una tesis vinculada a este proyecto. Así mismo servirá de base para la investigación y formulación de la tesis de Maestría en Administración de Negocios para el Ingeniero Diego Bolatti. Carrera de postgrado cursada en UTN Facultad Regional Resistencia.

Equipo de trabajo:

La estructura del equipo de trabajo es la siguiente:

Director	<ul style="list-style-type: none"> • Dr. Ing. Sergio Gramajo
Investigadores: (En orden alfabético)	<ul style="list-style-type: none"> • Ing. Diego Bolatti • Ing. Ricardo Calcagno • Ing. Carlos Cuevas • Ing. Reinaldo José Ramón Scappini
Colaborador Externo	<ul style="list-style-type: none"> • Dr. Luis Martínez López (Catedrático Universidad de Jaén, España) • Dra. Macarena Espinilla Estévez (Investigadora Universidad de Jaén)

- cision-making problems. *Computational Intelligence* 27, 489-512 (2011)
- Herrera, F., Herrera-Viedma, E.: Linguistic decision analysis: Steps for solving decision problems under linguistic information. *Fuzzy Sets and Systems* 115, 67-82 (2000)
 - Zadeh, L.: The concept of a linguistic variable and its application to approximate reasoning. Part I, Part II and Part III. *Information Sciences* 199-249, 301-357, 143-180 (1975)
 - <http://www.brianlinkletter.com/how-to-use-miniedit-mininets-graphical-user-interface/>
 - <https://www.sdnskills.com/learn/learn-python-1/>
 - <http://www3.fi.mdp.edu.ar/analisis/links/gnuplot-tut.pdf>
 - http://csie.nqu.edu.tw/smallko/sdn/iperf_mininet.htm

Referencias

- <https://www.opennetworking.org/sdn-definition/>
- <https://www.opennetworking.org/software-defined-standards/specifications/>
- <https://openflow.stanford.edu/display/ONL/POX+Wiki>
- <https://www.opendaylight.org/>
- <http://mininet.org/>
- <https://iperf.fr/>
- <http://gnuplot.info/>
- <https://www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/SDN-architecture-overview-1.0.pdf>
- Espinilla, M., Liu, J., Martínez, L.: An extended hierarchical linguistic model for de-

SERVIDOR VIRTUAL PARA DETECCIÓN DE INTRUSOS Y ATAQUES EN IPV6

D. E. Oneddu^(a); E. Zamudio^(b); N. B. Ganz^(b); M. J. Marinelli^(c)

(a) Facultad de Ciencias Exactas, Químicas y Naturales; Universidad Nacional de Misiones.

(b) Instituto de Materiales de Misiones; CONICET; Facultad de Ciencias Exactas, Químicas y Naturales; Universidad Nacional de Misiones.

(c) Departamento de Informática; Facultad de Ciencias Exactas, Químicas y Naturales; Universidad Nacional de Misiones.

oneddu@gmail.com

RESUMEN

La migración de IPv4 a IPv6 ha adquirido una importancia muy significativa, la cual cada día preocupa más a los proveedores de Internet, ya que no es posible simplemente dejar de utilizar el protocolo viejo y comenzar a usar el nuevo. Para solventar ésta necesidad de continuidad de negocio, surgen las llamadas técnicas o mecanismos de transición o coexistencia. Estas técnicas permiten la incorporación gradual de IPv6 dentro de una infraestructura IPv4 existente, ya que ambos protocolos no son compatibles por naturaleza. La seguridad de IPv4 e Ipv6 son totalmente independientes, por lo que se necesita establecer medidas para detectar, identificar, prevenir y tomar una decisión correcta frente a posibles ataques en ambos protocolos. Esto es una realidad que todos los usuarios de internet deben enfrentar en la actualidad y en particular los administradores de red.

El objetivo de esta línea de investigación es detectar de manera oportuna y a tiempo intrusos y/o ataques en un segmento de red IPv6 de una infraestructura en particular. Asimismo, se abordará puntualmente el IPv6, sus principales problemas al implementarlo y sus diferentes tipos de necesidades para configurarlo de manera segura.

Palabras Claves: IPv6, IDS, Servidor Virtual

CONTEXTO

Esta línea de investigación se lleva a cabo en el marco de una propuesta de Tesis de la Maestría en Tecnologías de la Información, de la Facultad de Ciencias Exactas, Químicas y Naturales (FCEQyN), de la Universidad

Nacional de Misiones (UNaM). Se proyecta este trabajo de investigación debido a la necesidad e importancia que tiene la ciberseguridad de IPv6 a nivel mundial, lo que nos motivó a investigar soluciones regionales o específicas para las diferentes unidades académicas de la universidad. Se consultó con los referentes informáticos de otras universidades nacionales tomando conocimiento de que si bien ocupan la tecnología IPv6, no tienen la necesidad de preocuparse por la configuración de su seguridad.

1. INTRODUCCIÓN

El agotamiento casi total de las direcciones IPv4 públicas en el 2011, llevó a grandes empresas a nivel mundial a participar del Día de IPv6 [1] en donde se puso a prueba si los servicios continuarían corriendo en dicho protocolo.

El IPv6 definido hace casi 20 años en el RFC 2460 de Diciembre de 1998 [2] prometía ser la solución al agotamiento de las direcciones publicas IPv4. Sin embargo, con el uso masivo actual que está generando IPv6, surgen ciertos inconvenientes no contemplados en la teoría y que son un verdadero problema al momento de utilizar este protocolo. Se podría enumerar algunos problemas de actual importancia en estos días:

En primer lugar, el problema más grande es que todavía se atraviesa la etapa de transición entre ambos protocolos. Esto implica, estar atados al uso ininterrumpido de IPv4 mientras se intenta evolucionar a IPv6. Lo cual infiere

una inversión económica y financiera importante, que en la mayoría de los países de Latinoamérica son muy difíciles de afrontar. Siendo necesario mantener equipos hardware y software funcionando en simultáneo con ambos protocolos. En el caso de los proveedores más importantes a nivel mundial, solamente los equipos de alta gama, es decir, los más costosos tienen la capacidad de realizar NAT64 [3]. En otras palabras, efectuar la traducción de un protocolo a otro. Razón por la cual es necesario un trabajo artesanal, para mantener funcionando satisfactoriamente toda la infraestructura de red y servidores en una organización, como lo es una Universidad Pública Nacional.

En segundo lugar, otro problema es que si bien IPv6 tiene la capacidad de identificar cada MAC address de cada dispositivo unívocamente a nivel mundial mediante la norma EUI64 [4], no existen suficientes Servidores DNS que resuelvan estas IP, o cada organización debe poseer su propio Servidor DNS; y provoca que una vez detectado el IP de un servidor, router o cualquier dispositivo crítico de una organización, éste sea propenso a ataques de nivel mundial.

En tercer lugar, otro problema es que IPsec definido dentro de la misma cabecera IPv6 no cumple satisfactoriamente su función de seguridad impenetrable. Además, al utilizar Stateless address autoconfiguration (SLAAC)¹ los sistemas operativos como Windows y Mac OS X, preferirán usar IPv6 en una red siempre que sea posible, lo que hereda el problema descrito en el punto anterior. IPv6 está ideado para autoconfigurarse por completo, por lo tanto, un atacante podría configurar un router

¹ Autoconfiguración de dirección sin estado es un método en el que a la interfaz del host o enrutador se le asigna un prefijo de 64 bits, y los últimos 64 bits de su dirección son derivados por el host o enrutador con la ayuda del proceso EUI-64 que se describe en las siguientes líneas. SLAAC usa el protocolo NDP para funcionar.

fraudulento y obtener automáticamente información de la red de una organización en particular, cuyos usuarios comenzarán a usar su servidor DNS falso.

Es importante destacar que actualmente las diferentes configuraciones hechas en IPv6 no contemplan adecuadamente diferentes tipos y niveles para la seguridad y protección ante ataques de DoS² y DDoS³ con lo cual es importante, además de tener un buen firewall correctamente configurado, algún sistema de detección de intrusos y posteriormente también un sistema de prevención de intrusos. En este trabajo de investigación se propone desarrollar una alternativa para abordar algunos aspectos de los problemas descritos, mediante la detección de intrusos y ataques, en el ámbito de redes de datos institucionales como lo son las redes de las Universidades Públicas de la República Argentina. Implicaría la utilización de los equipos ya existentes. Sin tener que adquirir equipamiento nuevo y costoso, como sería la solución de los grandes proveedores. Esta solución no debe afectar al desempeño normal de los equipos ni comprometer su seguridad. Será necesario para esto la configuración de un servidor virtual que detecte ataques en IPv6 mediante un sistema de detección de intrusos (IDS); y que debe trabajar en conjunto con un router de borde IPv6-Only que no interfiera con la arquitectura actual de la red institucional IPv4-Only en producción. En ningún momento y bajo ninguna circunstancia se tomará la información contenida en el área de datos de los paquetes, en primer lugar por considerarse irrelevante

² Denegación de Servicio (DoS) es una interrupción en el acceso de un usuario autorizado a una red informática, generalmente una causada con intención maliciosa

³ Un ataque de denegación de servicio distribuido (DDoS) es un intento de hacer que un servicio en línea no esté disponible al abrumarlo con tráfico de múltiples fuentes. Apuntan a una gran variedad de recursos importantes, presentan un gran desafío para asegurarse de que las personas puedan publicar y acceder a información importante.

para esta investigación, y en segundo lugar por cuestiones de la ley de Protección de Datos Personales⁴.

Con lo referente a la legislación, la asignación de dirección IPv6 y las regulaciones que poseen, la Universidad Nacional de Misiones ya tiene su pool de direcciones IPv6 asignados por LACNIC, del cual cada unidad académica cuenta con una porción de dicho pool para uso según sea necesario [5]–[7].

Todas las pruebas en la red de la unidad académica se realizarán sin sugestionar a los usuarios finales de dicha red, es decir, no dar aviso que se está implementando un nueva tecnología, o seguridad de protocolo nuevo en este caso; de esa manera se podrá evaluar la transparencia que tiene el estudio y la notificación que tienen los usuarios con respecto al uso cotidiano de la red. Por otra parte, una vez que esté todo configurado y funcionando, se harán simulaciones de ataques programados y aleatorios, con lo cual se solicitará la ayuda y participación de algunos administradores de red de otras unidades académicas; con el objetivo de obtener retroalimentación, sugerencias y optimización de configuraciones en el estudio de caso.

Toda la investigación será abordada en relación al modelo OSI y no al TCP/IP, es decir siempre que se mencione los niveles o capas será tenido en cuenta en relación al Modelo OSI, a continuación se expone en la siguiente imagen una comparación entre ambos modelos para su mejor comprensión.



Figura 1: Comparación del Modelo OSI y TCP/IP

[8], [9]

Lo admirable de IPv6 es su encabezado, una dirección IPv6 es 4 veces más grande que una IPv4, pero sorprendentemente, el encabezado de una dirección IPv6 es solo 2 veces más grande que el de IPv4. Los encabezados IPv6 tienen un encabezado fijo y cero o más encabezados opcionales (extensión). Toda la información necesaria que es esencial para un enrutador se guarda en el Encabezado fijo. El encabezado de extensión contiene información opcional que ayuda a los enrutadores a entender cómo manejar un paquete / flujo. En la Figura 2 se puede apreciar la estructura del encabezado IPv6.

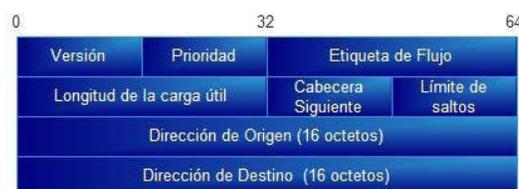


Figura 2: Cabecera IPv6

[8], [9]

Como hemos mencionado anteriormente, la UNaM posee su propio pool de direcciones IPv6 asignado por LACNIC, este pool se segmenta para el uso de las diferentes unidades académicas de dicha universidad. En la siguiente imagen se aprecia el enmascaramiento IPv6 de la UNaM.

⁴ Ley 25.326: Disposiciones Generales. Principios generales relativos a la protección de datos. Derechos de los titulares de datos. Usuarios y responsables de archivos, registros y bancos de datos. Control. Sanciones. Acción de protección de los datos personales.

Sancionada: Octubre 4 de 2000.

Promulgada Parcialmente: Octubre 30 de 2000.

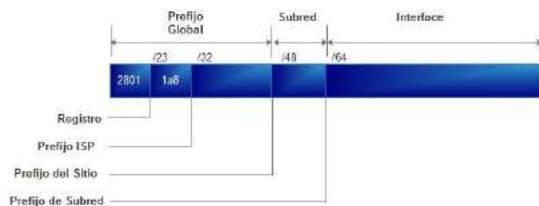


Figura 3: Máscaras IPv6

[10]

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo tiene como principal línea de investigación el tratamiento de la seguridad en el contexto de redes IPv6 mediante la virtualización de servidores en datacenter en Universidades Nacionales. Instalando y configurando servicios de detección de intrusos y ataques; así como también la utilización de Maquetas para la enseñanza de ciberseguridad en redes y comunicaciones. También se podrá anexar o incorporar placas SBC como dispositivo router, bridge, firewall, switching, etc. con el mismo fin pedagógico.

3. RESULTADOS ESPERADOS

Se pretende obtener un servidor virtual correctamente configurado con las necesidades básicas para la detección de intrusos y ataques en IPv6, totalmente migrable y exportable a otras instituciones o unidades académicas de las diferentes universidades nacionales. Un producto software multiplataforma totalmente independiente de la infraestructura a la cual se quiere aplicar.

Con respecto a las enseñanzas de redes y comunicaciones se pretende desarrollar maquetas para diferentes trabajos prácticos que ayuden a los alumnos a comprender el complejo mundo de las comunicaciones y conocer cuales son los elementos a tener en cuenta en la utilización de ciberseguridad

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto es parte de un plan de tesis de maestría en Tecnologías de la Información, en el cual se prevén incorporar becarios, tesisas, ayudantes de cátedra para Sistemas Operativos, Comunicación y Redes I y Comunicación y Redes II de las Carreras de Analista en Sistemas de Computación, Licenciatura en Sistemas de Información y Profesorado Universitario en Computación, de la Facultad de Ciencias Exactas, Químicas y Naturales (FCEQyN), de la Universidad Nacional de Misiones (UNaM).

5. BIBLIOGRAFÍA

- [1] Ole J. Jacobsen, "The Internet Protocol Journal," *Cisco Syst.*, vol. 171, no. 5, p. A20, 2011.
- [2] S. Deering and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," *RFC 2460*, 1998. [Online]. Available: <https://www.ietf.org/rfc/rfc2460.txt>. [Accessed: 10-Jul-2017].
- [3] G. Tsirtsis and P. Srisuresh, "Network Address Translation - Protocol Translation (NAT-PT)," *RFC 2766*, 2000. [Online]. Available: <https://www.ietf.org/rfc/rfc2766.txt>. [Accessed: 07-Feb-2016].
- [4] J. Abley, "Resource Records for EUI-48 and EUI-64 Addresses in the DNS," *RFC 7043*, 2013. [Online]. Available: <https://tools.ietf.org/html/rfc7043>. [Accessed: 13-Jul-2017].
- [5] "IPv6 Subredes y DNS UNaM," 2016. [Online]. Available: <http://www.tcpiputils.com/browse/ipv6-address/2801:1a8::-2801:1a8:ff:ffff:ffff:ffff:ffff:ffff>. [Accessed: 06-Feb-2016].
- [6] IANA, "IPv6 Global Unicast Address Assignments," 2016. [Online]. Available: <http://www.iana.org/assignments/ipv6-unicast-address-assignments/ipv6-unicast-address-assignments.xhtml>. [Accessed: 05-Feb-2016].

- [7] “Asociación Redes de Interconexión Univeritaria.” [Online]. Available: <http://www.riu.edu.ar/topologia.html>. [Accessed: 27-Feb-2016].
- [8] A. S. Tanenbaum, *Redes de computadoras*. Pearson Educación, 2003.
- [9] W. Stallings, *Comunicaciones y Redes de Computadores*, Séptima Ed. Pearson Educación, 2004.
- [10] R. Graziani, “CiscoPress - IPv6 Fundamentals: A Straightforward Approach to Understanding IPv6,” *IPv6 Fundam.*, p. 456, 2013.

Puerta de Enlace para Internet de las Cosas usando Computadora Industrial Abierta.

Carlos Taffernaberry, Gustavo Mercado, Matías Pecchia,
Sebastian Tobar, Ariel Verdejo, Juan Sayago

GridTICs – Grupo en Tecnologías de la Información y las Comunicaciones

Departamento de Electrónica / Facultad Regional Mendoza / UTN

{carlos.taffernaberry, gustavo.mercado, matias.pecchia,
sebastian.tobar, ariel.verdejo, juan.sayago}@gridtics.frm.utn.edu.ar

Resumen

Este trabajo describe un prototipo de software y hardware libre, que fue empleado exitosamente como gateway (pasarela) en sistemas de Internet de las Cosas (IoT - Internet of Things).

Normalmente un gateway de IoT es el dispositivo que actúa de interfaz de conexión entre dispositivos internos a la red IoT y el mundo exterior, normalmente la Internet.

El prototipo fue construido tomando como base de la placa CIAA (Computadora Industrial Abierta Argentina) a la que se adicionaron hardware/software de comunicación interna y externa.

La comunicación interna se basó en una pila de protocolos. El protocolo de capa de enlace usado fue el protocolo IEEE 802.15.4, que es el que normalmente utilizan las Redes de Sensores Inalámbricos. Los protocolos de capas superiores fueron el conjunto de la familia TCP/IP, entre los que tiene significativa importancia el “middleware” 6LoWPAN. Este último permitió que todo el sistema de IoT (interno y externo) pueda utilizar el nuevo protocolo de red denominado IPv6 de manera transparente y extremo a extremo, con las ventajas que esto representa.

El sistema operativo de IoT denominado “Contiki” fue usado en los dispositivos de la Red de Sensores Inalámbricos.

En el gateway fue implementado el SO FreeOSEK.

El almacenamiento de los datos sensados y transmitidos por los motes se hizo en una base de datos del tipo Time-Series llamada InfluxDB y la representación de los mismos

fue realizada usando la plataforma abierta Grafana.

Palabras Clave: Internet of Things, Internet de las Cosas, 6LoWPAN, Computadora Industrial Abierta Argentina (CIAA), OpenMote.

Contexto

El presente trabajo está inserto en el proyecto de investigación acreditado por la Universidad Tecnológica Nacional código EIUTIME0003646TC denominado “RED SIPIA-LP: Estudio de mecanismos de bajo consumo energético para aplicar a una red de sensores inalámbricos en el ámbito de agricultura de precisión”. El proyecto fue llevado adelante por investigadores y becarios de la Facultad Regional Mendoza.

Uno de los objetivos de este proyecto fue “Estudiar, simular y evaluar los protocolos de comunicación de internet de las cosas”. Para cumplir con ello se participó y se obtuvo financiamiento de la convocatoria “Proyectos de Innovación, Desarrollo y Adopción de la Computadora Industrial Abierta Argentina (CIAA) del MINCYT, de acuerdo a resolución Ministerial 613/15.

El trabajo en la convocatoria se denominó “GW-CIAA-IoT: Gateway con CIAA para red inalámbrica de IoT” y es el detallado en el presente trabajo.

1. INTRODUCCIÓN

Gracias a los avances y reducción de costos en dispositivos electrónicos y de comunicación inalámbrica, es posible, en la actualidad, construir sistemas de sensores

multifuncionales y multipropósito de bajo costo que operan con poca energía, de un tamaño pequeño, y con una capacidad de comunicación del tipo inalámbrica a corta distancia.

Estos dispositivos, denominados comúnmente motes o motas, constan de una unidad de procesamiento con un poder de cómputo mínimo, memoria, un módulo de comunicación inalámbrica y uno o varios dispositivos de sensado que capturan parámetros como temperatura, aceleración, humedad, etc. Un conjunto de motes comunicados entre sí es lo que se conoce como una Red de Sensores Inalámbrica (WSN - Wireless Sensor Network) [1].

OpenMote [2] es un fabricante de motes de hardware abierto para el desarrollo de la Internet de las Cosas. Estos dispositivos son complementados por un conjunto de herramientas de software de código abierto. La plataforma de OpenMote fue diseñada “a medida” para correr con el sistema operativo de código abierto OpenWSN [3] usando los estándares de IoT. También es posible el uso de OpenMote con otros sistemas operativos para embebidos de código abierto tales como FreeRTOS, RIOT and Contiki-OS. La combinación seleccionada de OpenMote y Contiki-OS nos permitió desarrollar, para el presente trabajo, las características de bajo consumo energético [4] y [5].

Por otro lado, hace 20 años, Internet se usaba principalmente como herramienta para acceder a información. En los últimos 10 años hemos vivido una nueva forma de uso de Internet, donde todo se ha convertido en social, transaccional y móvil. En la actualidad estamos atravesando una nueva transformación, en la que cada objeto tiene una identidad virtual propia y es capaz de integrarse e interactuar de manera independiente con cualquier otro objeto, sistema, o humanos. De esta forma estamos en presencia de un nuevo cambio en nuestra forma de vida, creándose nuevos modelos de negocio, productos y compañías, denominándose a esto la Internet de las Cosas (IoT - Internet of Things). Este desarrollo no

sería posible sin el soporte del nuevo protocolo de red de Internet, llamado IPv6, que permite contar con direcciones de red suficientes como para dotar a cada uno de los componentes de una red de sensores de una dirección pública IPv6.

Puntualmente, el ámbito de las WSN ha despertado un interés especial para lograr asociar a cada uno de los motes que la componen una WSN a la IoT. Contemplando los requerimientos de la IoT y las WSNs de área personal (PAN), un grupo de trabajo de la IETF desarrolló el protocolo 6LoWPAN [6] que brinda soporte a redes LowPAN para el protocolo IPv6.

Por medio de esta capa de adaptación 6LoWPAN, como se puede observar en la Fig. 1, se permite la interacción de los nodos de sensores que usan el protocolo 802.15.4, con el backbone de Internet conectado al protocolo 802.3, permitiendo que cada sensor pueda ser accedido desde cualquier host en cualquier parte del mundo.

Para permitir extender el alcance de una WSN, debido a que no siempre todos los nodos están en alcance entre sí, se creó otro grupo de trabajo en la IETF llamado ROLL (Routing Over Low power and Lossy Networks) para evaluar problemas de encaminamiento y proponer soluciones. El resultado fue el diseño de un protocolo de enrutamiento para LLN (Low power and Lossy Networks) denominado RPL [7], con soporte de una variedad de capas de enlace (IEEE 802.15.4, IEEE 802.15.4g o Powerline Communication), que comparten características comunes, como ser bajo ancho de banda, alta tasa de pérdidas y baja potencia.

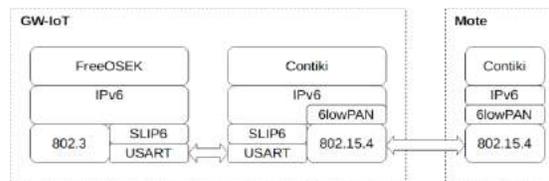


Fig 1. Capas de Protocolos de IoT

El Proyecto CIAA (Computadora Industrial Abierta Argentina) [8] nació en el año 2013 como una iniciativa conjunta entre el sector

académico y el industrial con el objetivo de impulsar el desarrollo tecnológico nacional, darle visibilidad positiva a la electrónica argentina y generar cambios estructurales en la forma en que se desarrollan y utilizan los conocimientos. Todo esto en el marco de un trabajo libre, colaborativo y articulado entre la industria y la academia.

La CIAA es la primera y única computadora del mundo que reúne dos cualidades:

- Es industrial: ya que su diseño está preparado para las exigencias de confiabilidad, temperatura, vibraciones, ruido electromagnético, tensiones, cortocircuitos, etc. Que demandan los productos y procesos industriales
- Es abierta: ya que toda la información sobre su diseño de hardware, firmware, software, etc. está libremente disponible en internet bajo la Licencia BSD, para que cualquiera la utilice como quiera. Por otro lado, su diseño no está atado a los procesadores de una determinada compañía, como ocurre con otras computadoras abiertas.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Como fue indicado en el apartado Contexto, el **Objetivo Principal** para el diseño e implementación del Gateway fue:

“Diseñar, ensayar e implementar una mejora de un sistema de redes inalámbricas agrícolas aplicadas, usando tecnología CIAA.”

Objetivos secundarios

- Análisis y estudio de protocolos de comunicación de la IoT aplicado a WSN.
- Diseño y desarrollo de hardware de Gateway de red SIPIA [9] de WSN utilizando CIAA.
- Evaluación y selección de sistema operativo embebido aplicable a CIAA con soporte de protocolos IoT.
- Diseño y desarrollo de software de Gateway de red SIPIA de WSN utilizando CIAA.
- Evaluación, selección e implementación de motes de WSN con soporte de IoT.

- Capacitación y difusión de los conocimientos adquiridos, como compromiso social en el ámbito empresarial y académico.

3. RESULTADOS OBTENIDOS

El objetivo general, de acuerdo a lo planteado en [10], se cumplió en su totalidad. Se logró implementar un sistema sobre la plataforma CIAA para comunicaciones extremo a extremo desde Internet hacia redes Inalámbricas de Sensores. El sistema utilizó el protocolo IPv6 en CIAA para acceder a Internet, haciendo uso del firmware oficial del proyecto CIAA. Adicionalmente se utilizó hardware de OpenMote Technologies para implementar la red Inalámbrica de Sensores.

En cuanto a los objetivos secundarios planteados, se puede decir que:

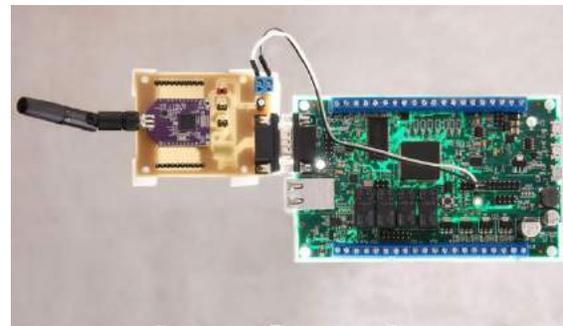


Fig 2: Gateway compuesto por placa CIAA, placa de interconexión y placa OpenMote.

- Fueron estudiadas distintas alternativas de protocolos de comunicación y se optó por continuar con la propuesta original de 6LoWPAN por ser éstos estándares abiertos del IETF.
- Se diseñó y desarrolló un circuito impreso que interconecta una CIAA-NXP con un mote OpenMote CC2538 o compatible, para permitir la implementación de SLIP6 [11] entre ambos, de acuerdo a la Fig 2.
- Se estudió la posibilidad de usar distintos sistemas operativos para la plataforma CIAA, entre ellos Linux, Mbed OS, RiotOS, FreeRTOS y FreeOSEK usado en el proyecto CIAA. Linux fue descartado porque el procesador de la CIAA-NXP no cuenta con Unidad de Gestión de Memoria. Mbed OS está orientado para embebidos y dispositivos ARM, pero no está portado para LPC4337 (el

Sistema en Chip Principal de la plataforma CIAA) y limita el desarrollo a esos dispositivos. Riot no presentaba documentación ni mucha actividad como proyecto al momento de estudio. Descartando las opciones anteriores, se consideró que el proyecto era viable usando FreeRTOS [12] o FreeOSEK [13]. Finalmente fue seleccionado para hacer el gateway el firmware oficial del proyecto CIAA (CIAA-Firmware), que está basado en FreeOSEK. De esta manera se propuso hacer nuevos aportes al proyecto CIAA y así generar vinculación local.

- Debido a que el firmware del proyecto CIAA no constaba con soporte para protocolos IoT, el trabajo implicó el diseño e implementación de los protocolos IPv6 y SLIP6. Este diseño y el código correspondiente está descrito en [14].

- Fueron seleccionado el proyecto los motes OpenMote CC2538 de OpenMote Technologies por ser compatibles con firmware Contiki, Openwsn, Riot entre otros asequibles y tener como componente principal el SoC (Sistema en un Chip) Texas Instruments CC2538 que está presente también en otros dispositivos que proveen distintos fabricantes de hardware. Adicionalmente, OpenMote fue la opción más conveniente, respecto al costo-beneficio del producto.

- El software adicional de almacenamiento de la información fue desarrollado sobre una plataforma de PC compatible, con Sistema Operativo GNU-LINUX y lenguaje de programación Python. Esta aplicación es un servidor de socket tipo datagrama, que recibe los datos que envían los motes, y los almacena en una Base de Datos INFLUXdb [15], ideal para series temporales, especialmente diseñada para IoT.

- La presentación de la información se llevó a cabo mediante una plataforma abierta para monitoreo y análisis de información llamada Grafana [16], la cual puede ser accedida por un cliente HTTP o Browser desde cualquier lugar del mundo con conectividad a Internet. Fue instalado sobre una plataforma de PC compatible, con Sistema Operativo GNU-

LINUX.

En la Fig 3 se puede observar una captura de pantalla de los datos graficados,



Fig 3: Información de gráficos presentados por Grafana

4. FORMACIÓN DE RECURSOS HUMANOS

Uno de los principales objetivos del proyecto es la capacitación de los recursos humanos.

La meta como investigadores es fortalecer la capacidad para realizar investigación científica, generar conocimientos y facilitar la transferencia de tecnología que permita el desarrollo humano.

Este proyecto de investigación posibilitará la colaboración inter-institucional y la ejecución de proyectos conjunto entre grupos I+D de diferentes disciplinas.

Para lograr estos objetivos se dispuso del siguiente personal:

- 2 Investigadores formados
- 1 Becario doctoral (beca UTN)
- 1 Becario graduado (Beca BINID UTN)
- 2 Becarios alumnos (Beca alumno UTN)
- 1 Tesista de carrera de grado

Adicionalmente se realizaron:

- Dictado de Cursos, Seminarios y Conferencia para público especializado.
- Promoción, coordinación y asistencia técnica de tesis de grado para alumnos de la FRMza.
- Promoción, coordinación y asistencia técnica a pasantes alumnos, cursantes de carreras de grado y de pre-grado en el ámbito de la UTN FRMza.

- Promoción, coordinación, dirección y asistencia técnica a Tesis doctorales, postgrado y/o maestría.
- Presentación de Trabajos en Congresos y Reuniones Técnicas/Científicas.
- Publicación de Trabajos en revistas con/sin referato.
- Publicación de todo el código fuente desarrollado en el presente proyecto, en <https://github.com/GridTICs/gw-ciaa-iot>.

5. BIBLIOGRAFÍA

- [1] Peter Waher, "Learning Internet of Things", 1st Edition, O'Reilly Ed.
- [2] OpenMote: Open-Source Prototyping Platform for the Industrial IoT Conference Paper · September 2015 ISSN 1867-8211
- [3] "OpenWSN" disponible en <http://www.openwsn.org/> consultado el 15 de Marzo de 2018.
- [4] A. Moschitta, and I. Neri. "Power consumption assessment in wireless sensor networks." ICT-energy-concepts towards zero-power information and communication technology. InTech, 2014.
- [5] The ContikiMAC Radio Duty Cycling Protocol. Adam Dunkels adam@sics.se. SICS Technical Report T2011:13. ISSN 1100-3154. December 2011.
- [6] S. Chakrabarti, "IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN)", RFC 8066, ISSN: 2070-1721, IETF, February 2017
- [7] T. Winter, Ed. "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, ISSN: 2070-1721, IETF, March 2012
- [8] Proyecto CIAA, Disponible en: <http://www.proyecto-ciaa.com.ar/> consultado el 14 Marzo 2018
- [9] A. Diedrichs, C. Taffernaberry, G. Mercado, G. Grunwaldt, M. Pecchia, G. Tabacchi, M. González, N. Altamiranda, "RED SIPIA-LP Estudio de mecanismos de bajo consumo energético aplicados a Red de Sensores Inalámbricos en el ámbito de Agricultura de Precisión", WICC 2016, ISBN: 978-950-698-377-2, Abril 2016.
- [10] C. Taffernaberry, G. Mercado, "GW-CIAA-IoT: Gateway con CIAA para red inalámbrica de IoT", WICC 2016, ISBN: 978-950-698-377-2, Abril 2016
- [11] J. Romkey, A "Non Standard For Transmission Of IP Datagrams Over Serial Lines: SLIP", RFC 1055, June 1988
- [12] FreeRTOS. Disponible en <https://www.freertos.org/>, consultado el 14 de Marzo de 2018.
- [13] FreeOSEK. Disponible en: <http://opensek.sourceforge.net/> consultado el 14 de Marzo de 2018.
- [14] Código abierto del proyecto disponible en <https://github.com/GridTICs/gw-ciaa-iot/>
- [15] "InfluxDB open-source time series database". Disponible en: <https://www.influxdata.com> consultado el 14 Marzo 2018.
- [16] Grafana "The open platform for beautiful analytics and monitoring". Disponible en: <https://grafana.com> consultado el 16 de Marzo de 2018.

PREDICCIÓN Y ALERTA TEMPRANA DE INCENDIOS FORESTALES MEDIANTE INTEGRACIÓN DE WSN E IMÁGENES SATELITALES EN UN MÉTODO DE REDUCCIÓN DE INCERTIDUMBRE GUIADO POR DATOS

Méndez-Garabetti Miguel^{1,2,3}, Bianchini Germán², Caymes-Scutari Paola^{2,3},
Elgueta Rodrigo¹, Caylá Ivana Belén¹, Córdoba Diego¹, Peñasco Andrés¹, Morelli, José¹,
Carrizo Marcelo¹, Varela Pablo¹, Giorlando Agustín¹, Lujan Arnaldo¹ y Carballo Nicolás¹

¹Universidad de Mendoza, Facultad de Ingeniería (UM)

²Laboratorio de Investigación en Cómputo Paralelo/Distribuido

Departamento de Ingeniería en Sistemas de Información

Facultad Regional Mendoza/UTN

³Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

miguel.mendez@um.edu.ar, gbianchini@frm.utn.edu.ar, pcaymesscutari@frm.utn.edu.ar,
rodrigo.elgueta@um.edu.ar, iva.cayla@gmail.com, diego.cordoba@um.edu.ar,
penasco.andres@gmail.com, jose.morelli@um.edu.ar, carrizomarcelo@gmail.com,
jazxz.pv@gmail.com, agus.giorlando@gmail.com, aarrnaall@gmail.com,
nicolas.carballo@gmail.com

RESUMEN

Los incendios forestales son unos de los fenómenos naturales que más perjuicios causan en nuestro país y en el mundo, generando pérdidas de vidas humanas, daños en la propiedad, la flora, la fauna, el suelo, entre otros. Debido a esto, se considera de interés desarrollar herramientas, sistemas, métodos, estrategias, entre otros, que permitan minimizar los efectos negativos causados por éstos. El presente proyecto pretende realizar un aporte significativo en esta área, dado que propone integrar información de diferentes fuentes, en un método paralelo de reducción de incertidumbre guiado por datos aplicado a la predicción y alerta temprana de incendios forestales. Las fuentes de información estarán conformadas por: redes de sensores inalámbricos, imágenes satelitales, históricos de incendios forestales, entre otros. Dicha integración permitirá determinar la ocurrencia precoz de los incendios forestales antes de que estos se

magnifiquen, y además permitirá reducir la incertidumbre de algunas de las variables dinámicas que alimentan al modelo de predicción de comportamiento de incendio, impactando de forma positiva en la calidad de predicción del sistema.

Palabras clave: alerta temprana, predicción de incendios, incendios forestales, WSN, sensores, imágenes satelitales

CONTEXTO

El presente proyecto de I+D cuenta con el financiamiento de la Dirección de Investigaciones de la Universidad de Mendoza (DIUM), el mismo se desarrolla en la Facultad de Ingeniería de la Universidad de Mendoza con la colaboración del Laboratorio de Investigación en Cómputo Paralelo Distribuido de la Universidad Tecnológica Nacional, Facultad Regional Mendoza (UTN-FRM).

1. INTRODUCCIÓN

Los incendios forestales causan grandes daños y pérdidas [1], en Mendoza, estos eventos ocurren principalmente entre los meses de septiembre y marzo. En 2016 según el informe de Estadísticas de Incendios Forestales desarrollado por el Ministerio de Ambiente y Desarrollo Sustentable [2], en Argentina los incendios devastaron un total de 1.072.642 hectáreas, y puntualmente en Mendoza, la cifra alcanzó un total de 93.529 hectáreas.

Con el propósito de reducir la incidencia de los incendios forestales y mitigar sus efectos negativos, se vuelve necesario contar con herramientas que permitan:

- a) **Detectar en el menor tiempo posible la ocurrencia de estos fenómenos:** este tipo de herramientas también se conocen como sistemas de alerta temprana, los cuales cumplen un rol muy importante en el proceso de lucha contra incendios debido a que permiten actuar antes que el incendio se extienda demasiado, posibilitando así la utilización de menor cantidad de recursos en el proceso de extinción y, sobre todo, generando menor cantidad de daños [3,4,5].
- b) **Pronosticar con antelación el comportamiento de la línea de fuego:** Cuando los incendios ya se han magnificado, predecir su comportamiento puede ser la herramienta que marque la diferencia. La predicción del comportamiento de un incendio forestal puede permitir gestionar los generalmente “escasos” recursos disponibles, de manera eficiente y segura, facilitando la toma de decisiones complejas tales como evacuaciones, planificación de cortafuegos, extracción de combustible, entre otros [6,7,8,9].

El presente proyecto aborda el diseño de una WSN para ser aplicada como sistema de alerta temprana de incendios forestales, además pretende integrar de forma colaborativa los datos adquiridos mediante dicha red junto a los generados, para tal fin, por las diferentes agencias espaciales (imágenes satelitales). Dicha información será utilizada para alimentar un método de predicción del comportamiento de incendios forestales que se beneficiará de las diferentes escalas de trabajo de cada plataforma tecnológica. En base a lo dicho, la fundamentación teórica la podemos dividir en tres secciones: a) redes de sensores inalámbricas aplicadas a la detección de incendios forestales, b) utilización de productos satelitales para la detección de incendios forestales, y c) métodos de reducción de incertidumbre aplicados a la predicción del comportamiento de incendios forestales.

Redes de sensores inalámbricos (WSNs):

El uso de redes de sensores que cubren grandes territorios asegurando el monitoreo efectivo de fenómenos tales como incendios forestales sigue siendo un problema de particular interés y significación en la comunidad científica. Esto es debido a que las WSN ofrecen soluciones económicas y eficientes de monitoreo, dado que estos dispositivos son baratos y compactos, pudiendo ser desplegados en extensos territorios y al funcionar conjuntamente y de forma autónoma pueden utilizarse eficazmente para detectar gases peligrosos y detectar los incendios forestales [10].

Una WSN es un conjunto de elementos autónomos, interconectados de manera inalámbrica, que colaboran con el objetivo de resolver una tarea en común. Los sensores toman del medio la información y la convierten en señales eléctricas, posteriormente los nodos, los cuales contienen varios sensores toman los datos de los sensores y los envían a las estaciones base. La comunicación entre los nodos suele realizarse mediante el estándar IEEE 802.15.4,

utilizando el stack de protocolos ZigBee, aunque también puede utilizarse 802.11 o Bluetooth.

Desde el punto de vista de la detección de incendios, existen diferentes alternativas para detectar la ocurrencia de éste fenómeno, en el ámbito de las WSN la metodología más usada consiste en medir la temperatura ambiente, la humedad, la luz y la presión barométrica, aunque también puede detectarse mediante el sentido de humo, CO/H₂ y gases peligrosos.

Productos satelitales para la detección de incendios forestales:

Debido a su cobertura repetitiva de gran área, los datos de satélite son útiles para la detección de incendios, monitoreo y evaluación del área quemada en tiempo casi real. Los focos de calor MODIS son bastante utilizados como datos de entrada para la generación de eventos de incendios. La detección de focos de calor se basa en la capacidad de un sensor remoto de capturar la energía emitida por la superficie terrestre en las longitudes de onda correspondientes al infrarrojo medio y térmico. De acuerdo a las temperaturas emitidas por los incendios (entre 300°C y 1500°C), se establecen umbrales de temperatura que permiten detectar los frentes de fuego activos. El proceso de detección es influenciado por una diversidad de factores, que comprenden al tipo de vegetación quemada, el nivel de combustible acumulado, las condiciones ambientales, y cuestiones relativas a la geometría de observación de la plataforma satelital. Cada foco que se detecta activo (en llama) al momento del registro de la imagen, es representado sobre la superficie terrestre mediante la coordenada central del pixel. El tamaño de pixel MODIS correspondiente a los canales térmicos es de 1000 mts de lado, lo cual representa una superficie de 10.000 mts cuadrados (1 ha.). La CONAE procesa focos de calor a partir de la versión original del algoritmo MOD14. A lo largo de los últimos años, el algoritmo MOD14 ha pasado por diversas pruebas y validaciones que

han resultado en mejoras sobre su capacidad de detección. La colección 5 (C5) es el producto de focos de calor actual, el cual, provee mejoras en cuanto a la calidad de detección de incendios pequeños. Además, también existen adaptaciones de este algoritmo para poder incorporar datos de otras plataformas, como GOES y NOAA-AVHRR, y en particular los provistos por el sensor VIIRS, el cual corresponde a la nueva generación de sensores que a futuro continuará con el monitoreo de incendios en la era post MODIS. Es importante mencionar que las imágenes satelitales de estas misiones se encuentran accesibles a toda la comunidad por lo que pueden ser procesadas para obtener la información necesaria.

Métodos de reducción de incertidumbre:

Debido a la imprecisión de los parámetros de entrada y la dificultad de medirlos en tiempo real, es que ha sido necesario recurrir a técnicas que sean capaces de reducir la incertidumbre, como por ejemplo los Métodos Guiados por Datos (DDM, Data Driven Methods). Los DDM consideran un gran número de valores para cada parámetro, posteriormente, realizan una búsqueda (mediante alguna técnica de optimización heurística) para encontrar un conjunto de parámetros que describa, de la mejor manera posible, el comportamiento anterior del fuego el cual se espera que pueda ser utilizado para predecir el siguiente instante de tiempo.

En otras palabras, los DDM realizan una calibración para obtener estos valores "óptimos" de los parámetros de entrada. Sin embargo, estos métodos obtienen un solo conjunto de valores, y para aquellos parámetros que poseen un comportamiento dinámico, el valor encontrado no es generalmente útil para describir correctamente el comportamiento del modelo. Esta categoría se denomina Métodos Guiados por Datos de Solución Única [11,12,13]. Existe otra clasificación de los DDM los cuales operan con casos solapados y combinaciones de parámetros para efectuar las predicciones. Esta categoría se denomina Méto-

dos Guiados por Datos con Múltiples Soluciones Solapadas (DDM-MOS, Data Driven Methos with Multiple Overlapping Solutions).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La detección temprana de los incendios forestales permite evitar la magnificación de estos fenómenos, permitiendo tomar decisiones antes de que estos alcancen a cubrir grandes superficies. De esta manera se puede permitir minimizar los efectos negativos que éstos generan.

Por tal motivo la presente línea de investigación consiste en el desarrollo de una Red Inalámbrica de Sensores (Wireless Sensor Network, WSN) [14,15] para ser aplicada como sistema de alerta temprana de incendios forestales, además pretende integrar de forma colaborativa los datos adquiridos mediante dicha red junto a los generados, para tal fin, por las diferentes agencias espaciales (imágenes satelitales). Dicha información será utilizada para alimentar un método de predicción del comportamiento de incendios forestales que se beneficiará de las diferentes escalas de trabajo de cada plataforma tecnológica. Tales desarrollos permitirán ofrecer una herramienta de alerta temprana de incendios forestales y mejorar la capacidad de predicción los métodos actuales.

3. RESULTADOS OBTENIDOS/ESPERADOS

Tras el desarrollo del presente proyecto se espera:

1. Implementar una red de sensores inalámbricos que opere como sistema de alerta temprana.
2. Identificar variables sensadas que puedan ser incorporadas al método de predicción, de esta manera se podrá reducir la incertidumbre en ciertas variables que alimentan al modelo.

3. Integrar el sistema de senado al sistema de predicción del comportamiento de incendios forestales.

4. FORMACIÓN DE RECURSOS HUMANOS

El presente proyecto está a cargo del Mg. Ing. Miguel Méndez-Garabetti y el Dr. Germán Bianchini, además se cuenta con el asesoramiento -como profesional invitado- de la Dra. Paola Caymes-Scutari. La formación de recursos humanos comprende tanto a estudiantes de grado (Ingeniería en Informática, Ingeniería en Electrónica, Bioingeniería) y de posgrado (Maestría en Teleinformática) todos de la Universidad de Mendoza.

También se cuenta con la participación de investigadores en formación quienes se desempeñan como docentes de grado de en la Universidad de Mendoza: Ing. Diego Córdoba e Ing. Rodrigo Elgueta. Este último se encuentra realizando su tesis de maestría en relación a temáticas centrales del presente proyecto.

5. BIBLIOGRAFÍA

- [1] Morgan, P., Hardy, C., Swetnam, T.W., Rollins, M.G., & Long, D.G. (2001). Mapping fire regimes across time and space: Understanding coarse and fine-scale fire patterns, *International Journal of Wildland Fire.*, 10, 329-342.
- [2] Estadísticas de Incendios Forestales 2016. (2016).
- [3] Yemail Lemaitre, A. J. & Moreno Ramos, A. I. Sistema para alerta temprana de incendios forestales. (2016).
- [4] Lic Nicolás Mari, P. A. & Scavuzzo Fabiano Morelli, M. Diseño de un Sistema de Alerta y Respuesta Temprana a Incendios de Vegetación – SARTiv. (2012).
- [5] SPARC (Organization), M. & Universidad Tecnológica de Pereira. *Scientia et technica. Scientia Et Technica* **18**,

- (Universidad Tecnológica de Pereira, 1995).
- [6] Bianchini, G., Cortés, A., Margalef, T., Luque Fadón, E., Chuvieco, E., & Camia, A. (2005). Wildland fire risk maps using S2F2M. *JCS&T*, 5, no. 4, 244–249
- [7] Bianchini, G., Caymes-Scutari, P., & Méndez-Garabetti, M. (2015) Evolutionary-statistical system: a parallel method for improving forest fire spread prediction, *J. Comput. Sci.* 6 (2015) 58–66.
- [8] Altintas, I., Block, J., de Callafon, R., Crawl, D., Cowart, C., Gupta, A., & Smarr, L. (2015). Towards an Integrated Cyberinfrastructure for Scalable Data-driven Monitoring, Dynamic Prediction and Resilience of Wildfires. *Procedia Computer Science*, 51, 1633–1642.
- [9] Méndez-Garabetti M., Bianchini G., Caymes-Scutari P., & Tardivo M. (2016) Increase in the quality of the prediction of a computational wildfire behavior method through the improvement of the internal metaheuristic. *Fire Safety Journal*, 82:49–62. ISSN 0379-7112.
- [10] Antoine-Santoni, T., Santucci, J. F., de Gentili, E. & Costa, B. Using Wireless Sensor Network for Wildfire detection. A discrete event approach of environmental monitoring tool. in *2006 First International Symposium on Environment Identities and Mediterranean Area* 115–120 (IEEE, 2006).
doi:10.1109/ISEIMA.2006.344927
- [11] Beven, K., & Binley, A. (1992) The future of distributed models: model calibration and uncertainty prediction, *Hydrol. Process.* 6 279–298.
- [12] Piñol, J., Salvador, R., & Beven, K. (2002) Model Calibration and Uncertainty Prediction of Fire Spread, *Forest Fire Research & Wildland Fire Safety*, on CD-ROM, Millpress
- [13] Abdalhaq, B. (2004) A methodology to enhance the prediction of forest fire propagation (Ph.D. thesis), Universitat Autnoma de Barcelona, Spain.
- [14] Baronti, P., Pillai, P., Chook, V. W. C., Chessa, S., Gotta, A., & Hu, Y. F. (2007). Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards. *Computer Communications*, 30(7), 1655–1695.
- [15] Leblon, B. (2005). Monitoring Forest Fire Danger with Remote Sensing. *NH*, 35(3), 343–359.

CIAA EN LA INDUSTRIA MISIONERA: FABRICA DE LADRILLOS. INTEGRACIÓN CON SOFTWARE SCADA Y ERP

Juan de Dios Benítez^c, Ruben Syniuk, Cristian Cabral, Diego Alberto Godoy^a, Edgardo A. Belloni^b, Eduardo O. Sosa^c, Hernán Bareiro^f, Fabián Favret^g

Centro de Investigación en Tecnologías de la Información y Comunicaciones
(C.I.T.I.C.) Departamento de Ingeniería y Ciencias de la Producción-Universidad
Gastón Dachary
Av. López y Planes 6519- Posadas, Misiones, Argentina. Teléfono: +54-376-4438677

^adiegodoy@citic.ugd.edu.ar, ^bebelloni@ugd.edu.ar, ^ceduardo.sosa@citic.edu.ar,
^dfidelis.sergio@citic.edu.ar, ^ejuan.benitez@citic.ugd.edu.ar, ^fhbareiro@citic.ugd.edu.ar,
^gfabianfavret@citic.ugd.edu.ar

Resumen

En este trabajo se presenta un proyecto de investigación denominado “Simulación en las Tics: Diseño de Simuladores de Procesos de Desarrollo de Software Agiles y Redes De Sensores Inalámbricos para la Industria y la Academia”, cuyo objetivo es diseñar simuladores de procesos de desarrollo de software agiles y de redes de Sensores Inalámbricos para la Industria y la academia. En este artículo se trabaja sobre la automatización de una fabrica mediante el uso de sistemas embebidos [9], precisamente la adaptación de un FIRMWARE opensource, CIAA FIRMWARE[3], añadiendole características nuevas para poder monitorear y gestionar procesos productivos de manera remota por medio de un software SCADA[1] y un ERP[2].

Palabras claves: SCADA; ERP; Remoto; FIRMWARE; OpenSource; Proyectos de Desarrollo de Software; .

Contexto

El trabajo presentado en este artículo tiene como contexto marco el proyecto de investigación denominado “Simulación en las Tics: Diseño de Simuladores de Procesos de Desarrollo de Software Agiles y Redes De Sensores Inalámbricos para la Industria y la Academia”, registrado actualmente en la Secretaría de Investigación y Desarrollo de la Universidad Gastón Dachary (UGD) con el número Código IP A07003 y radicado en el Centro de Investigación en Tecnologías de la Información y Comunicaciones de dicha universidad.

El mismo fue incorporado como proyecto aprobado en el llamado a presentación interna de la UGD de proyectos de investigación N°7 mediante la Resolución Rectoral 07/A/17 y es una continuidad del Proyecto Simulación como herramienta para la mejora de los procesos de software desarrollados con metodologías ágiles utilizando dinámica de sistemas, R.R. UGD N° 18/A/14 y R.R. UGD N° 24/A/15.

Entre las líneas con mayores resultados dentro del proyecto referido, se encuentran las de: Construcción de una plataforma de gestión y simulación de

datos de redes de sensores inalámbricos, una interfaz web para el simulador de WSN Shawn, Sistemas de gestión de residuos de la ciudad de posadas con tecnologías de Internet de la cosas, Sistema de monitoreo de la temperatura en el proceso de secado del Té.

Introducción

La evolución natural en la mejora de los procesos de una empresa manufacturera que posee actividades de producción manuales y semi-automáticas es una migración de estas hacia la automatización total con las herramientas disponibles más actuales. La mejora de las actividades de los procesos industriales pueden ser parciales o totales de acuerdo a los aspectos y sectores que estuvieran involucrados. En el caso que aquí tratamos se considera un sistema de control de un sector de planta del proceso de fabricación, su supervisión y su habilitación para el soporte de programación de la producción. La supervisión en un sistema de control es una actividad fundamental en un proyecto de automatización industrial. Los componentes que se disponen en un sistema de este tipo para lograr el nivel de supervisión varían de acuerdo a los requerimientos y restricciones del proyecto. En algunos casos es suficiente con la disposición de tableros con indicadores luminosos y sonoros que indican los diferentes estados del proceso. En otros casos se desarrollan interfaces HMI (Human Machine Interface) con pantallas digitales que muestran en gráficos de calidad los estados y etapas del proceso de manera muy comprensible. También y a partir de las tecnologías de aplicaciones Web,

se dispone de interfaces HMI desarrolladas como clientes ligeros para dispositivos de escritorio como también para dispositivos móviles.

La programación de la producción es una actividad de gran importancia en las plantas de producción en las que el volumen y frecuencia de tareas es alta. Actualmente se cuenta con herramientas de software que facilitan las tareas de programación. También es posible que los sistemas de control automático del proceso de planta y estas herramientas trabajen de manera coordinada por un mecanismo automático de compartición de datos. En este proyecto de automatización industrial basado en una industria de fabricación de ladrillos, se requiere desarrollar e implementar una aplicación de control basada en la arquitectura de hardware y software que presenta la CIAA[3][4], la comunicación entre ésta, un sistema SCADA (Supervisory Control And Data Acquisition) [8] y un sistema ERP (Enterprise Resource Planning)[5]. Estos componentes de software son fundamentales para proveer la funcionalidad requerida para la administración del proceso de la planta. La aplicación de control se desarrollará con el firmware del proyecto CIAA[10]. Su función incluye el accionamiento de motores, revisión de sensores, registro del estado del proceso y recepción de parámetros de configuración. Además provee soporte de comunicación para otras etapas de control y otros sistemas. Para el monitoreo y supervisión del proceso se emplea un sistema SCADA[8], que permite realizar un seguimiento en línea del estado de las variables. Para la programación de la

producción se utiliza un sistema ERP[5], que permite relacionar las variables del proceso con registros necesarios para la programación de órdenes de fabricación.

Línea de Investigación

En esta línea de investigación se han planteado varios objetivos. El objetivo principal de la misma corresponde a: “Proponer una capa de integración, basada en servicios para el sistema operativo de tiempo real OSEK[3] del CIAA FIRMWARE, independiente de la aplicación de control para el soporte de comunicación con SCADA y ERP” donde para poder lograr dicho objetivo se plantearán objetivos particulares y secundarios los cuales se detallan a continuación:

- Comprender el funcionamiento de interfaces de integración utilizadas en sistemas embebidos para RTOS[4].
- Analizar y proponer especificaciones para una capa de integración basada en servicios independiente de la aplicación de control para el CIAA Firmware.
- Modelar y desarrollar la capa de integración basada en servicios independiente de la aplicación de control para integración de SCADA con ERP.
- Verificar el correcto funcionamiento de la capa de integración desarrollada con el proyecto tomado como piloto.

Resultados

El proyecto aquí presentado consta de la automatización de un sistema de transferencia de material pesado en un ambiente hostil como es el de una fabrica de ladrillos en la provincia de Misiones. El progreso y desarrollo del mismo se divide en dos puntos primordiales. 1.- Control del carro de transferencia de material dentro de la fabrica. Automatizacion del traslado. 2.- Integración del sistema de transferencia a un software de control de procesos SCADA y un software de programación de tareas ERP. Donde en esta etapa contaremos con una división interna que se sentra en la resolucion de la problemática de la conexión de interfaces de comunicación y la de una Propuesta de arquitectura que contemple todo lo anterior. La primera etapa de control del sistema de transporte dentro de la fabrica se analiza y se desarrolla el software correspondiente para el control del carro de transferencia (Fig. 1)



Fig.1: Fotografía del carro

En la fig. 2 se puede ver un diagrama de cómo se mueve el carro dentro de la fabrica.

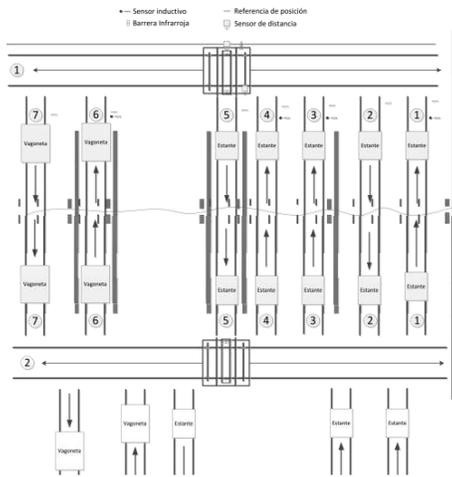


Fig.2: Movimiento del carro

Para resolver la problemática de las comunicaciones, teniendo en cuenta el tipo de ambiente y distancias de operatividad, se decide trabajar sobre arquitecturas de comunicaciones inalámbricas, en este caso el desarrollo se basará sobre la norma de comunicación de WLAN 802.11 a través del módulo ESP8266.

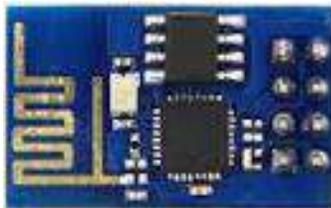


Fig.3: Placa de comunicación inalámbrica

Ya pasando a la última etapa del sistema tenemos la presentación de una arquitectura de trabajo para poder llevar a cabo el proceso completo. Donde la misma debería contemplar la siguiente estructura (Fig.4).

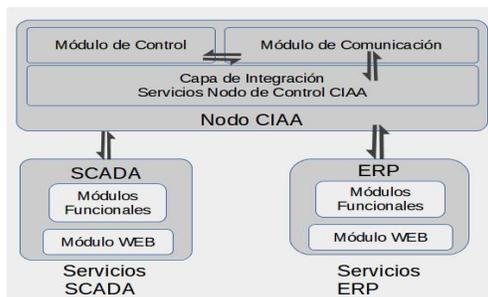


Fig.4: Arquitectura propuesta

Esta arquitectura implementada sobre la Fig. 2 donde se ve el movimiento del carro quedaría de la siguiente manera Fig.5.

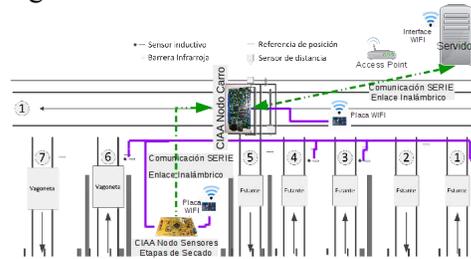


Fig.5: Implementación de la arquitectura

De la Fig. 4 y Fig.5 se aprecia que aparece un servidor. En este servidor se encontrarán alojados los sistemas SCADA[8] y ERP[5].

La complejidad de este proyecto radica en 2 puntos. El primero, el manejo del Firmware[5] base de la CIAA y del sistema operativo de tiempo real que la misma utiliza, ya que como dicha computadora se encuentra en etapas de desarrollo no existe bibliografía y documentación de donde basarse para los desarrollos. Y basados en eso la segunda problemática radica en la dificultad de la integración del CIAA Firmware con los softwares tanto SCADA como ERP.

Formación de Recursos Humanos

El equipo de trabajo se encuentra formado por cuatro investigadores con distintos niveles de posgrado, un Doctor en Ciencias Informáticas y Magister en Redes de Datos; un Doctor en Tecnologías de la Información y Comunicaciones Magister y Especialista en Ingeniería de Software; un Maestrando de Ingeniería de la Web; dos Maestrando en Redes de Datos y ocho estudiantes en período de realización de trabajos finales de grado en el contexto de las carreras de Licenciatura en Sistemas de

Información y de Ingeniería en Informática de la UGD. Actualmente, el número de tesinas de grado aprobadas en el contexto de este proyecto, es de cinco, y otras tres en proceso de desarrollo. El número de tesis de maestría terminadas relacionadas con este proyecto es de una.

Bibliografía

- [1] A. R. Penin, *Sistemas SCADA*, 2° Edición. Marcombo, 2007.
- [2] L. M. González, *ERP: guía práctica para la selección e implantación*, 1° Edición. Grupo Planeta (GBS), 2004.
- [3] M. Cerdeiro, *Breve introducción a OSEK-VDXUn sistema operativo de tiempo real estandarizado*, 1° Edición. Proyecto CIAA, 2015.
- [4] C. M. Sendis, *OSEK/RTOS & OSEK turbo Introduction*, 1° Ed. NXP:Freescale, 2009.
- [5] J. W. Valvano, *Embedded Microcomputer Systems: Real Time Interfacing*, 3° Edición. Cengage Learning, 2011.
- [6] IEEE. *eSOA – Service Oriented Architectures adapted for Embedded Networks*. IEEE, 2009. 7ma INDIN 2009.
- [7] Jim Keogh. *J2EE Manual de Referencia*. Mc Graw Hill, 1° edición edition, 2003.
- [8] Aquilino Rodríguez Penin. *Sistemas SCADA*. Marcombo, 2° edición edition, 2007.
- [9] Publicación de la SASE. ¿qué son los sistemas embebidos? *Simposio Argentino de Sistemas Embebidos*, 2010.
- [10] Christian Michel Sendis. *OSEK/RTOS OSEK turbo Introduction*, 1° ed. edition, 2009.
- [11] Richard Zurawski. *Industrial Communication Technology Handbook*. CRC Press, 2° edición edition, 2014.

Rendimiento de Aplicaciones Web en plataformas de contenedores de código abierto

Lic. Gabriel Antonio Coronel, Mgter Oscar Adolfo Vallejos
Universidad de la Cuenca del Plata
email: gabriel.c190@gmail.com; oval_lejos@hotmail.com

Abstract

Cloud computing makes extensive use of virtual machines, because they allow the workload is isolated and also, the use of resources is somewhat controlled. Currently, the additional levels of abstraction involved in virtualization allow to reduce workload performance. The most recent advances in container-based virtualization simplify the deployment of software applications without this implying the loss of control over the resources assigned to different applications. In this paper, we analyze the performance of a web application running in the containers of the Docker platform, as well as that of another free software platform, KVM, and a host operating system.

Resumen

La computación en la nube hace un uso extensivo de máquinas virtuales, pues permiten que la carga de trabajo esté aislada y además, el uso de los recursos este algo controlado. En la actualidad, los niveles adicionales de abstracción involucrados en virtualización permiten reducir el rendimiento de carga de trabajo. Los avances más recientes en la virtualización basada en contenedores simplifican la implementación de aplicaciones de software sin que esto implique la pérdida el control sobre los recursos asignados a las diferentes aplicaciones. En este documento, se analiza el rendimiento de una aplicación web funcionando en los contenedores de la plataforma Docker¹, como también el de otra plataforma de software libre, KVM, y un sistema operativo anfitrión.

Palabras claves: máquinas virtuales, virtualización basada en contenedores, contenedores de código abierto.

Introducción

En la actualidad, en el mundo de las tecnologías de información, hay ciertas problemáticas que

afectan a diversas empresas y organizaciones en cuanto a la tecnología utilizada para su plataforma de computación en la nube. Además de poseer el equipamiento para ofrecer procesamiento y almacenamiento suficiente para cubrir los requerimientos del negocio, son necesarios los medios para otorgar recursos de forma eficaz y eficiente cuando el cliente los requiera. Con la virtualización los servidores estarán preparados para responder a cambios en las cargas de trabajo y al tener cada aplicación aislada con su propio servidor virtual se puede evitar que una aplicación impacte otras aplicaciones al momento de realizar mejoras o cambios.

Aún así, es posible encontrar ciertas deficiencias a la hora de implementar virtualización en una infraestructura. Entre las mismas se encuentran la sobrecarga de ejecutar múltiples máquinas virtuales, la dificultad de la administración de dependencias de las aplicaciones, problemas no presentes en entornos de desarrollo de software pero presentes en entornos de producción.

Estos problemas pueden solucionarse mediante el empleo de los contenedores de aplicaciones, que utilizan una virtualización de sistema operativo, pero surge la problemática del desempeño que pueda acarrear esta tecnología, de forma que existe la posibilidad de convertirse en una mala decisión implementar esta plataforma.

Docker es una forma de encerrar los servicios en entornos aislados, llamados contenedores, de modo que puedan ser empaquetados con todo lo que necesitan en términos de las bibliotecas y dependencias y el desarrollador puede estar seguro de que el servicio se ejecutará donde sea que Docker funcione [1]. Docker trae varias consideraciones que en las tecnologías anteriores no se advierten. Entre ellas, es que los contenedores son más fáciles de implementar y utilizar que los métodos anteriores, más allá de la ventaja de la asociación de Docker con otras empresas de contenedores, tales como Canonical, Google, Red Hat y Parallels, sobre su componente clave libcontainer de código abierto, que ha traído la estandarización, algo muy necesario en los contenedores.

Considerando que se trata de una tecnología nueva en constante cambio, no hay un conjunto

de buenas prácticas establecidas para que las empresas puedan migrar fácilmente a esta tecnología en estos días, sobre todo si no disponen de recursos como las empresas anteriormente mencionadas. Otra cuestión reside en la falta de documentación de rigor científico para sostener las ventajas de rendimiento que brindan los contenedores libres funcionando con la plataforma Docker.

Teniendo en cuenta lo anterior se realizó este trabajo en el que se llega a una conclusión de manera empírica acerca de las ventajas de performance de Docker comparado con una máquina virtual de QEMU-KVM y un sistema operativo anfitrión. KVM simplemente convierte el kernel de Linux en un hipervisor cuando se instala el módulo kernel de KVM... Para las emulaciones de entrada y salida, KVM utiliza un software de usuario, QEMU; este es un programa que hace emulación de hardware[2]. En este trabajo se seleccionaron una serie de métricas que se tuvieron en cuenta para realizar pruebas de rendimiento de una aplicación web funcionando con las tecnologías de virtualización mencionadas anteriormente, utilizando el software para medir el rendimiento de una aplicación web (rails perf test) y la herramienta de monitoreo de sistema de Ubuntu Linux, que proporciona datos precisos del sistema operativo anfitrión.

Método

Hoy en día, en un mundo cada vez más competitivo, el manejo de información se ha convertido en un factor determinante para el negocio de las empresas. Una buena elección de las tecnologías para su plataforma de sistemas de IT (tecnologías de información) posibilitará que el negocio tenga más posibilidades de asegurar una posición exitosa en el mercado.

En este contexto, las empresas inmersas en la computación en la nube, que se dedican a implementar software como servicio, dependen del rendimiento de su aplicación web para que su negocio prospere.

Si no se mide, el juicio puede basarse solamente en la evaluación subjetiva. Para permanecer en el mercado las empresas necesitan datos del desempeño de su servicio para poder hacer evaluaciones objetivas, por lo tanto, resulta muy importante la medición del software. El *IEEE Standard Glossary of Software Engineering*

Terminology [3] define métrica como "una medida cuantitativa del grado en el que un sistema, componente o proceso posee un atributo determinado". Las métricas de software proporcionan a los gerentes de desarrollo de software y administradores de proyectos una visión sensata de lo bien que sus aplicaciones funcionan en realidad. Aisladas, las métricas son sólo representaciones gráficas o visuales de datos. Sin embargo, mediante el análisis de las métricas, los desarrolladores de software pueden descubrir patrones, tendencias y problemas que afectan al rendimiento de las aplicaciones.

En los últimos años muchas empresas han implementado la virtualización de software para aportar mayor rendimiento a sus aplicaciones. La virtualización permite que una sola computadora contenga varias máquinas virtuales, cada una de las cuales puede llegar a ejecutar un sistema operativo distinto [4]. Siguiendo este camino aparece la plataforma de contenedores de Docker que utiliza una virtualización con menor sobrecarga, a diferencia de las otras máquinas virtuales.

Con esto, un equipo de desarrolladores puede llevar a cabo una aplicación como si utilizaran la misma máquina, logrando un mejor proceso para manejar las dependencias y facilitando el traslado del software al servidor u otras máquinas.

Una migración a contenedores podría facilitar el proceso de despliegue de una aplicación web. Con lo cual se reduciría el tiempo que le cuesta a una organización realizar el despliegue de su software. Al mismo tiempo resultaría más simple el trabajo de mantener entornos de desarrollo con diferentes sistemas operativos de forma estable. Según el servicio de monitoreo Datadog[5] en las empresas que adoptan Docker, los contenedores tienen una vida útil media de 2,5 días, mientras que en todas las empresas, las máquinas virtuales tradicionales en la nube tienen una vida media de 23 días. Lo cual indica la simplicidad para subir una nueva instancia de una aplicación en un contenedor y reemplazar otra comparado con las VM.

Este trabajo provee una comparación actualizada de una aplicación web con el uso de contenedores, máquina virtual y sistema operativo huésped utilizando hardware reciente y software con una serie de pruebas para analizar sus respectivos rendimientos.

Como objetivo general se evaluó las ventajas y desventajas con métricas de rendimiento para una aplicación web utilizando un sistema operativo huésped, un sistema operativo virtualizado y un contenedor.

Para ello fue necesario: a) Desarrollar una aplicación web que funcione en la plataforma de contenedores Docker, en un sistema operativo virtualizado y sistema operativo huésped; b) Seleccionar, aplicar y analizar las métricas de rendimiento comparando el uso sobre el sistema operativo, un sistema operativo virtualizado y un contenedor; c) Analizar las ventajas y desventajas relacionadas con el manejo de dependencias.

Se realizaron las siguientes etapas:

- Se realizó una revisión sobre las tecnologías para poder desarrollar una aplicación web, es decir, los frameworks web que se utilizan actualmente en el desarrollo de software ágil, teniendo como resultado la definición de qué tecnologías utilizar a la hora de llevar a cabo el desarrollo.
- Se llevó a cabo el desarrollo del software funcionando con los contenedores de Docker. Este software luego será ejecutado también en una máquina virtual y un sistema operativo huésped.
- Se realizó una búsqueda de las métricas de software enfocadas en el rendimiento y se seleccionaron las más adecuadas.
- Posteriormente, se aplicaron las métricas para contrastar entre estos diferentes entornos, con lo cual se obtuvieron los resultados, que fueron analizados.
- Finalmente, se elaboraron las conclusiones correspondientes.

Desarrollo de la experiencia

Se seleccionaron los diez primeros frameworks web más utilizados en la actualidad, considerando aquellos que incluyan la programación del lado del servidor de acuerdo a la web hotframework[6].

De esta selección, se optó por la utilización del framework Ruby on Rails versión 4.2, debido a su interfaz amigable y a la concreción de un prototipo en un periodo corto de tiempo, con lo cual resulta muy productivo.

La aplicación web consiste en una librería en línea, es decir, permite comprar libros por internet de una librería para luego ser enviados al cliente. Dicha aplicación sirve como un caso para poder ejecutar aplicación usando un sistema operativo anfitrión, un contenedor de Docker y una máquina virtual para poder obtener las métricas relacionadas al rendimiento. El desarrollo fue llevado a cabo utilizando una metodología híbrida de extreme programming. Como lenguaje se utilizó una implementación de Ruby en Java, lo que permite una mejor concurrencia con aplicaciones exigentes.

Docker es multiplataforma y para poder ejecutarse en Windows y Mac OS debe utilizarse un hipervisor que ejecute una máquina virtual de Linux en el cual estará en funcionamiento del contenedor, por lo tanto la manera más eficiente para esto es utilizar un sistema operativo Linux. Se seleccionó para esto el sistema Ubuntu por ser un sistema amigable y con mucha documentación al respecto.

Durante la ingeniería de requerimiento se utilizaron el diagrama de flujo de datos, diagramas de casos de uso y diagrama de clases, que fueron refinados en la etapa del diseño.

Después de que se hayan hecho los diagramas, no se inició directamente la codificación, sino que se desarrolló una serie de pruebas unitarias y una prueba de integración, para cumplir con esa prueba se llevó a cabo un desarrollo iterativo e incremental. Para estas pruebas se utilizó la herramienta que viene con el framework, llamada MiniTest, la cual es una mejora de la herramienta original de testing de Rails Test::Unit. Se eligió este software al ser el recomendado por quienes intervienen en el desarrollo del framework Ruby on Rails y tener una sintaxis sencilla para escribir los tests. Una vez que el código estuvo terminado, se le aplicó de inmediato una prueba unitaria, con lo que se obtiene retroalimentación instantánea.

Para poder utilizar un contenedor y ejecutar la aplicación web en un dicho es necesario realizar un Dockerfile que contenga las instrucciones necesarias para crear una imagen de sistema operativo y posteriormente un contenedor (ver **Figura 1**). Para poder generar una imagen realmente pequeña para ejecutar la aplicación web, se utilizó la versión 17 de Docker (Edición Comunitaria) y 3.6 de Alpine Linux. Esta distribución incluye sólo los archivos mínimos necesarios para arrancar y ejecutar el sistema operativo.

Posterior a eso se incluye la configuración para instalar Ruby on Rails y las librerías necesarias. El último paso consiste en ejecutar la instrucción que construya la imagen y ejecuta el contenedor.

```
FROM jruby:9.1.7-alpine
MAINTAINER myself@gabriel@gmail.com
ENV BUILD_PACKAGES="curl-dev ruby-dev build-base" \
    DEV_PACKAGES="lib-dev libxml2-dev libssl-dev tzdata yaml-dev sqlite-dev" \
    RUBY_PACKAGES="ruby ruby-libs console ruby-json yaml nodejs" \
    RAILS_VERSION="4.2.5"
RUN \
  apk --update --upgrade add $BUILD_PACKAGES $RUBY_PACKAGES $DEV_PACKAGES $S \
  gen install -N bundler
RUN gen install -N nokogiri -- --use-system-libraries $S \
  gen install -N rails --version "$RAILS_VERSION" $S \
  echo "gen: --no-documents" -- --j-gencrc $S \
  cp -f ./gencrc /etc/gencrc $S \
  chmod uog+r /etc/gencrc $S \
  # cleanup and settings
  bundle config --global build.nokogiri -- --use-system-libraries $S \
  bundle config --global build.nokogumbo -- --use-system-libraries $S \
  find / -type f -iname '*.age-new -delete $S \
  rm -rf /var/cache/apk/* $S \
  rm -rf /usr/lib/lib/ruby/gems*/cache/* $S \
  rm -rf -j-gen
RUN mkdir -p /usr/src/app
WORKDIR /usr/src/app
COPY . /usr/src/app
RUN bundle install
EXPOSE 3000
# The main command to run when the container starts. Also
# tell the Rails dev server to bind to all interfaces by
# default.
CMD ["bundle", "exec", "rails", "server", "-b", "0.0.0.0"]
```

Figura 1: Dockerfile utilizado para la aplicación

A continuación se seleccionaron métricas para la medición de las siguientes características:

- Medir el tiempo del mundo real transcurrido durante la prueba de funcionamiento (wall time). Diferencia entre el momento en el que termina una tarea y la hora a la que comenzó la tarea. Se ve afectada por cualquier otro proceso al mismo tiempo que se ejecutan en el sistema.
- Medir el tiempo del usuario. Cantidad de tiempo que la CPU gastó en modo de usuario, es decir, dentro del proceso. Esto no se ve afectado por otros procesos y por el tiempo que posiblemente se gasta bloqueado.
- Medir la cantidad de memoria utilizada para el caso de prueba de rendimiento. El uso de la memoria proporciona la información sobre la cantidad de espacio libre en disco.
- Medir la cantidad de tiempo gastado en recolección de basura para el caso de prueba de rendimiento.

Se realizaron varias pruebas diferentes, las cifras que se muestran son las mejores que se obtuvieron. Esto se pudo realizar mediante la herramienta rails perf test versión 0.0.7. Esta herramienta permite hacer un tipo especial de pruebas de integración, diseñada para probar el rendimiento de una aplicación web hecha con el

framework web Ruby on Rails. Se eligió este software al ser un proyecto con desarrollo activo y open source. Con esto se realizó una prueba de estrés para exigir al sistema los más posible. La prueba es similar a un test de integración, se carga la página inicial pero con una gran cantidad de productos y se realiza una compra. Cuando se utilizó las herramientas de virtualización se tuvo en cuenta el estado del sistema operativo anfitrión (consumo de CPU y memoria RAM) con la herramienta de monitoreo del sistema, la cual viene por defecto en el sistema operativo Ubuntu.

Resultados Obtenidos

Para el sistema operativo anfitrión se utilizó el sistema operativo Ubuntu 16.04 con el kernel 4.10, que luego sería el anfitrión de la máquina virtual y el contenedor, se realizó la prueba de rendimiento (ver **Tabla 1**).

Wall time	Tiempo de usuario	Memoria
124 mili segundos	57 mili segundos	7,1 megabytes

Tabla 1: Resultados sistema operativo

Para el sistema operativo invitado se utilizó una máquina virtual de Ubuntu creada con QEMU-KVM en su versión 2.5, administrando dicha máquina virtual con el programa virt-manager versión 1.3.2 e instalando la aplicación web en dicha máquina se realizó la prueba de rendimiento (ver **Tabla 2**). Utilizando el monitor de sistema de Ubuntu se aprecia que los CPU llegaron a un pico de 100%. El consumo de memoria de sistema operativo tuvo un máximo de 2,1 gigabytes y se obtuvieron las métricas correspondientes. En estado de reposo la máquina virtual consume 1,9 gb (gigabytes).

Wall time	Tiempo de usuario	Memoria
369 mili segundos	137 mili segundos	6,8 megabytes

Tabla 2: Resultados de máquina virtual

Midiendo el sistema sin exigirlo después de haberlo iniciado, se obtuvo estas cifras: 33% de uso de CPU y 1,9 gigabytes de memoria ram. Al momento de exigirlo llegó a un 100% de uso de CPU con 2,1 gigabytes de memoria.

Utilizando un contenedor Docker a partir de una imagen de sistema operativo personalizada para las dependencias de la aplicación web se realizó la prueba de rendimiento (ver **Tabla 3**). En este caso los CPU llegaron a un pico de 95% y 97%. El consumo de memoria de sistema operativo superó ligeramente al del hipervisor y se obtuvieron la

métricas correspondientes. En estado de reposo el contenedor consume 697,8 mb (megabytes).

Wall time		Tiempo de usuario	Memoria
133 segundos	mili	67 mili segundos	6,3 megabytes

Tabla 3: Resultados de Docker

Los resultados de la prueba son casi iguales que en el sistema operativo anfitrión.

Midiendo el sistema sin exigirlo después de haberlo iniciado, se obtuvo estas cifras: 0,13% de uso de CPU y 697,8 megabytes de memoria ram. Al momento de exigirlo llego a un 97% de uso de CPU con 2,2 gigabytes de memoria. En ese momento de exigencia el consumo de memoria del sistema operativo fue un poco mayor que en el caso de la máquina virtual.

Conclusión

Al finalizar este trabajo, se puede concluir que mientras QEMU-KVM es una solución implementada con mayor tiempo en desarrollo, teniendo en cuanto los resultados obtenidos con la herramienta de pruebas de rendimiento rails perf test, Docker es capaz de proveer un mejor rendimiento utilizando una menor cantidad de recursos, siendo la performance más cercana al de la aplicación funcionando sin virtualización, si bien al momento de ser exigidas ambas plataformas, con Docker el sistema operativo mostró una leve mayoría en uso de memoria. También se debe tener en cuenta que si bien hay alternativas

comerciales de hipervisores, en este caso QEMU-KVM es exclusivo para un sistema operativo Linux mientras que Docker cambió sus componentes para poder lograr contenedores en otras plataformas.

En el apartado de la administración de las dependencias para ejecutar la aplicación web, Docker tiene una manera más eficiente al tener el proceso automatizado declarando las dependencias necesarias en el Dockerfile mientras que en el caso de QEMU-KVM es un proceso manual por parte del usuario.

Llevar a cabo este proyecto, entendemos, contribuye disponer de un sostén científico respecto del rendimiento de contenedores de código abierto utilizados para virtualizar aplicaciones web. Con lo cual sería un factor importante en el momento que una organización decida migrar a esta tecnología.

Referencias Bibliográficas

- [1] Hane, O. (2015). *Build Your Own PaaS* with Docker. Birmingham, Reino Unido. Packt Publishing.
- [2] Mukhedkar P. (2016). *Mastering KVM Virtualization*. Birmingham, Reino Unido. Packt Publishing.
- [3] *IEEE Standard Glossary of Software Engineering Terminology*. (1993). IEEE.
- [4] Tanenbaum, A. (2009). *Sistemas Operativos Modernos*. Naucalpan de Juárez, México. Pearson Education.
- [5] (2017). 8 Surprising facts about real Docker adoption.
- [6] (2017). Hot Frameworks. Recuperado de: <https://hotframeworks.com/>

Bases de Datos y Minería de Datos

Modelo de Decisión para la Validación de Métodos de Imputación Mediante la Utilización de Algoritmos de Minería de Datos

Carlos R. Primorac¹, Julio C. Acosta^{1,2}, David L. La Red Martínez¹

¹ Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste
9 de Julio 1449, Corrientes (3400), Argentina

² Facultad de Ciencias Agrarias. Universidad Nacional del Nordeste
J. B. Cabral N° 2131, Corrientes (3400), Argentina

carlosprimorac@gmail.com, julioaforever@hotmail.com, lrmdavid@exa.unne.edu.ar

Resumen

Muchos de los conjuntos de datos (data sets) existentes u obtenidos en investigaciones científicas contienen valores faltantes (MVs: Missing Values) y anomalías (outliers) asociados a procedimientos de entrada manuales deficientes, mediciones incorrectas o errores en los instrumentos de medición. En minería de datos (DM: Data Mining) estas imperfecciones pueden afectar negativamente la calidad del proceso de aprendizaje supervisado o el rendimiento de algoritmos de agrupamiento de datos. La imputación es una técnica para reemplazar MVs con valores sustituidos. Pocos estudios informan una evaluación global de los métodos existentes con el fin de proporcionar directrices para hacer la elección metodológica más apropiada en la práctica. El propósito general de este trabajo es determinar un modelo de decisión que permita encontrar los métodos de imputación más adecuados para completar información faltante en un conjunto de datos mediante la utilización de algoritmos de DM.

Palabras Clave: valores faltantes, imputación, minería de datos, modelo de decisión.

Contexto

La propuesta se inserta dentro de una de las líneas de trabajo del Grupo de Sistemas Operativos y TICs (Res. 725/10 C.D. - FaCENA) en el marco del Proyecto de Investigación "Incidencia de los perfiles de los alumnos en el rendimiento académico en

Matemática del primer año de la Universidad", acreditado por la SGCyT - UNNE (PI: 16F002, Res. N° 970/16 C.S.).

Diversos estudios y publicaciones abordan la evaluación de rendimiento académico utilizando técnicas de DM [1] [2] [3] [4] [5].

En este proyecto de investigación se propone evaluar el rendimiento académico de los estudiantes en las asignaturas Algebra de la carrera Licenciatura en Sistemas de Información (LSI) de la Facultad de Ciencias Exactas y Naturales y Agrimensura (FaCENA) y Matemática I de la carrera Ingeniería Agronómica (IA) de la Facultad de Ciencias Agrarias (FCA) de la de la Universidad Nacional del Nordeste (UNNE) utilizando técnicas de DM.

Para definir los perfiles de los estudiantes y determinar patrones que conduzcan al éxito o fracaso académico, se implementará un modelo que relaciona las calificaciones de los estudiantes con otras variables, tales como factores socioeconómicos, demográficos, actitudinales, entre otros; en base a lo cual se clasificaran los diferentes perfiles de alumnos.

Los modelos predictivos buscados, permitirán tomar acciones tendientes a evitar el fracaso académico, detectando los alumnos con perfil de riesgo de fracaso académico de manera temprana, a poco del inicio del cursado de las asignaturas; lo que permitirá concentrar en ellos los esfuerzos de tutorías y apoyos especiales.

Introducción

Históricamente, la noción de descubrir patrones ocultos en los datos ha recibido una variedad de denominaciones incluidos el de DM y descubrimiento del conocimiento (KDD: Knowledge Discovery in Databases). KDD, se refiere al proceso general de descubrir conocimiento útil a partir de los datos. La DM es una etapa dentro del proceso general de KDD que se refiere a los medios algorítmicos mediante los cuales se extraen y enumeran patrones a partir de los datos [6].

Muchos de los conjuntos de datos existentes u obtenidos en investigaciones científicas contienen MVs y anomalías (outliers) asociados a procedimientos de entrada manuales deficientes, mediciones incorrectas o errores en los instrumentos de medición. La presencia de estas imperfecciones generalmente requiere de una etapa de preprocesamiento en la cual, con el fin de que resulten útiles y suficientemente claros para el proceso de extracción de conocimiento, los datos se deben preparar y limpiar [7] [8] [9] [10] [11] [12].

En DM se pueden encontrar tres problemas principales asociados con MVs y outliers: i) pérdida de eficiencia, ii) complicaciones en la manipulación y análisis de los datos y iii) sesgo resultado de las diferencias entre valores faltantes y completos [10] [12] [13]. Estos afectar negativamente la calidad del proceso de aprendizaje supervisado o el rendimiento del algoritmo de agrupamiento de datos [11].

En la literatura se proponen dos enfoques generales para enfrentarse a los MVs. En el caso más simple, las instancias con MVs se omiten. Una segunda alternativa es utilizar técnicas de imputación y estimarlos utilizando los datos existentes [8] [10] [14] [15] [16] [17].

Tradicionalmente, el tratamiento de los MVs se realizaba antes del análisis de los datos mediante métodos diseñados “ad hoc”.

Algunas de estas estrategias consistían en trabajar con información completa, eliminando todos los casos con MVs en una o más variables (listwise-deletion), considerando los casos con valores disponibles en la variable de análisis y descartándolos cuando contienen MVs (pairwise-deletion) o sustituyendo MVs con el promedio de la variable considerada. Sin embargo, el sesgo introducido por estas técnicas ha hecho que sean fuertemente criticadas en la literatura [18].

La imputación es una técnica para reemplazar MVs con valores sustituidos. Una característica importante para una instancia en particular puede imputarse [15]. Estos métodos utilizan diferentes algoritmos que se pueden dividir en imputación simple (single-imputation) e imputación múltiple (MI: Multiple Imputation) y, en los últimos años, se ha propuesto el uso de algoritmos de aprendizaje automático (ML: Machine Learning) [9] [10] [11] [19] [14] [16] [18] [20] [21].

La selección de un método de imputación depende del conjunto de datos, el mecanismo de pérdida de datos y los patrones, el porcentaje de MVs y el desempeño de la técnica de imputación utilizada [15].

El mecanismo de pérdida de datos es un factor clave para decidir el método de imputación a utilizar. Rubin [22] definió tres mecanismos por los cuales se genera la pérdida de datos: i) aleatoria (MAR: Missing at Random), completamente aleatoria (MCAR: Missing Completely at Random) y iii) no aleatoriamente (NMAR: Not Missing at Random).

El desempeño de un método de imputación, no solo depende de la cantidad de MVs, sino también de los patrones de pérdida. Existen diferentes patrones de MVs, algunos asociados con el registro y otros con el atributo. En el primer caso pueden ser simples, complejos, medios y mixto. En el segundo univariados, monótonos y arbitrarios [15]. Adicionalmente, el tamaño del conjunto

de datos y el porcentaje de MVs influye en la elección [16].

Diferentes técnicas de imputación funcionan bien sobre diferentes tipos de datos, algunas trabajan bien con enteros, otras únicamente con variables categóricas y algunas otras con datos combinados [12].

Finalmente, algoritmos de agrupamiento (no supervisados) y de clasificación (supervisados) se pueden adaptar para la imputación [14].

La mayoría de los artículos publicados en este campo se ocupan del desarrollo de nuevos métodos de imputación, sin embargo, pocos estudios informan una evaluación global de los métodos existentes con el fin de proporcionar directrices para hacer la elección metodológica más apropiada en la práctica [13].

1. Líneas de Investigación y Desarrollo

“Incidencia de los perfiles de los alumnos en el rendimiento académico en Matemática del primer año de la Universidad” es continuación de cuatro proyectos de investigación ejecutados desde el año 2004, oportunamente evaluados y acreditados en Comisión Externa; originados en la superpoblación y deserción de los alumnos en los cursos de trabajos prácticos de Álgebra en el primer año de la Universidad. Se trabajará en la búsqueda de las variables que inciden en el rendimiento académico de los alumnos, para ejecutar las acciones que permitan evitar la deserción, corrigiendo las situaciones detectadas que la generan.

2. Resultado Esperados

El propósito general de este trabajo es determinar un modelo de decisión que permita encontrar los métodos de imputación más adecuados para completar información faltante en un conjunto de datos mediante la utilización de algoritmos de DM.

Se espera poder determinar una metodología para seleccionar los métodos de imputación de datos más adecuados para imputar cada variable del conjunto de datos. Se utilizarán como métodos de validación de los métodos de imputación algoritmos de DM de eficacia reconocida. El criterio de validación que se utilizará será el de mayor similitud entre los resultados de los procesos de minería antes de la imputación (considerando solamente registros completos, excluyendo los registros con datos faltantes) y luego de la aplicación de cada uno de los métodos de imputación (incluyendo ahora los archivos completos, es decir los registros con datos ahora imputados), para lo cual habrá de definirse una métrica específica.

La metodología desarrollada será general, pero la aplicación de la misma será particular, para cada archivo con datos faltantes donde sea necesario imputar los mismos. La aplicación de la metodología general a desarrollar será parte de un modelo de decisión que se aplicará ante casos concretos de archivos con datos faltantes (la metodología se desarrollará como una herramienta en el modelo de decisión que incluye la especificación del contexto en el que se aplicará la metodología desarrollada).

Los objetivos específicos son:

- Definir la métrica a utilizar para la validación de los diferentes métodos de imputación aplicados al conjunto de datos.
- Definir el procedimiento de selección de los métodos de imputación más adecuados para ser aplicados al conjunto de datos.
- Definir el orden de prioridad de las variables para la aplicación de los diferentes métodos de imputación de datos en el conjunto de datos.

3. Formación de Recursos Humanos

El equipo de trabajo está compuesto por un Doctor, dos Magister y dos Licenciados en Sistemas de Información con cursados de

Maestría Finalizadas, de los cuales uno está desarrollando su Tesis en la propuesta presentada.

Referencias

- [1] D. L. la Red Martínez, J. C. Acosta, V. E. Uribe y A. R. Rambo, "Academic Performance: An Approach From Data Mining," *Journal of Systemics, Cybernetics and Informatics*, vol. 10, no. 1, pp. 66-72, 2012.
- [2] D. L. La Red Martínez, M. Karanik, M. Giovannini, N. Pinto, "Academic Performance Profiles: A Descriptive Model Based on Data Mining," *European Scientific Journal*, vol. 11, no. 9, pp. 17-38, March 2015.
- [3] D. L. La Red Martínez, G. Bobadilla Almada, "Estudio del rendimiento académico y detección temprana de perfiles de alumnos en la Facultad Politécnica de la Universidad Nacional del Este de Paraguay," en *Utilizando tecnologías en la educación para fortalecer la práctica docente en América Latina. Revisiones teóricas - Experiencias prácticas*, Bogotá, 2016.
- [4] D. L. La Red Martínez, M. Karanik, M. Giovannini, R. Scappini, "Towards to a Predictive Model of Academic Performance Using Data Mining in the UTN-FRRe," *Journal of Systemics, Cybernetics and Informatics*, vol. 14, no. 2, pp. 36-41, 2016.
- [5] D. L. La Red Martínez, M. E. Giovannini, M. E. Báez Molinas, J. I. Torre, N. Yaccuzzi, "Academic performance problems: A predictive data mining-based model," *Academia Journal of Educational Research*, vol. 5, no. 4, pp. 61-75. April 2017.
- [6] U. Fayyad, G. Piatetsky-Shapiro & P. Smyth, "From Data Mining to Knowledge Discovery in Databases," *AI Magazine*, vol. 17, no. 3, pp. 37-54, March 1996.
- [7] J.I. Peláez, J. M. Doña, D. L. La Red Martínez, "Fuzzy Imputation Method For Database Systems," in *Handbook of Research on Fuzzy Information Processing in Database*, Hershey, PA: IGI Global (701 E. Chocolate Avenue, Hershey, Pennsylvania, 17033, USA), 2008.
- [8] G. Madhu and T. V. Rajinikanth, "A novel index measure imputation algorithm for missing data values: A machine learning approach," in *2012 IEEE International Conference on Computational Intelligence and Computing Research*, Coimbatore, 2012, pp. 1-7.
- [9] J. Lengo, S. García & F. Herrera, "On the choice of the best imputation methods for missing values considering three groups of classification methods," *Knowledge and Information Systems*, vol. 32, issue 1, pp. 77-108, July 2012.
- [10] A. Farhangfar, L. A. Kurgan and W. Pedrycz, "A Novel Framework for Imputation of Missing Values in Databases," in *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 37, no. 5, pp. 692-709, Sept. 2007.
- [11] M. Pattanodom, N. Iam-On & T. Boongoen, "Clustering data with the presence of missing values by ensemble approach," in *2016 Second Asian Conference on Defence Technology (ACDT)*, Chiang Mai, 2016, pp. 151-156.
- [12] G. Rahman, Z. Islam, "A Decision Tree-based Missing Value Imputation Technique for Data Pre-processing," in *AusDM '11 Proceedings of the Ninth Australasian Data Mining Conference*, Ballarat, 2011, pp. 41-50.
- [13] P. Schmitt, J. Mandel & M. Guedj, "A Comparison of Six Methods for Missing Data Imputation," *Journal of Biometrics & Biostatistics*, vol. 6, issue 1, January 2015.
- [14] Y. Liu & V. Gopalakrishnan, "An Overview and Evaluation of Recent Machine Learning Imputation Methods Using Cardiac Imaging Data," *Data*, vol. 2, no. 1, p. 8, Jan. 2017.

[15] T. Aljuaid & S. Sasi, "Proper imputation techniques for missing values in data sets," in *2016 International Conference on Data Science and Engineering (ICDSE)*, Cochin, 2016, pp. 1-5

[16] B. Twala, M. Cartwright and M. Shepperd, "Comparison of various methods for handling incomplete data in software engineering databases," in *2005 International Symposium on Empirical Software Engineering*, 2005, pp. 105-114.

[17] Y. Vergouwe, P. Royston, K. G.M. Moons, D. G. Altman, "Development and validation of a prediction model with missing predictor data: a practical approach," *Journal of Clinical Epidemiology*, vol. 63, issue 2, 2010, pp. 205-214.

[18] James L. PeughCraig K. Enders, "Missing Data in Educational Research: A Review of Reporting Practices and Suggestions for Improvement," *Review of Educational Research*, vol. 74, issue 4, 2004, pp. 525 – 556.

[19] G. Tutz & S. Ramzan, "Improved methods for the imputation of missing data by nearest neighbor methods," *Computational Statistics & Data Analysis*, vol. 90, 2015, pp. 84-99.

[20] A. Pantanowitz & T. Marwala, "Evaluating the Impact of Missing Data Imputation through the use of the Random Forest Algorithm," <https://arxiv.org/abs/0812.2412>, Fecha de Acceso 18 de Diciembre de 2017.

[21] K. Arima, N. Okada, Y. Tsuji and K. Kiguchi, "Evaluations of a multiple SOMs method for estimating missing values," in *2014 IEEE/SICE International Symposium on System Integration*, Tokyo, 2014, pp. 796-801.

[22] D. B. Rubin, "Inference and Missing Data," *Biometrika*, vol. 63, issue 3, 1976, pp. 581-592.

Reconocimiento de Patrones Genéticos por Medio de Grafos

Cristóbal R. Santa María*, Romina Rebrij**Victoria Santa María***, Luis López* y
Marcelo Soria****

*DIIT-UNLaM, **Hospital Italiano (CABA), ***Instituto Lanari-FMed-
UBA, ****FAUBA Florencio Varela 1903 San Justo Pcia. de Buenos
Aires 54-011-44808952

csanta_maria@ing.unlam.edu.ar

rominarebrij@gmail.com

vctrstmr@gmail.com

llopez@ing.unlam.edu.ar

soria@agro.uba.ar

Resumen

Se expone la línea de investigación que se lleva adelante en el Departamento de Ingeniería e Investigaciones Tecnológicas de la UNLaM. Se detallan resultados del proyecto de investigación “Aplicación de Técnicas de Data Mining para Análisis del Microbioma Humano según Funcionalidades Metabólicas”, C200 del Programa de Incentivos. Con él se intenta aportar procedimientos para analizar la relación clínica entre el microbioma intestinal y la presencia de patologías. Esto comprende la obtención de muestras de microbiomas de pacientes, la identificación funcional de las secuencias genéticas y la determinación de la distribución de frecuencias por especies en cada paciente. En el proyecto de investigación anterior C169 se habían obtenido datos de secuencias del gen marcador 16S rRNA. La necesidad de establecer ahora una clasificación por funcionalidades metabólicas para todos los genes presentes en cada microbioma, llevó a la búsqueda de nuevos datos crudos (no ensamblados) y al análisis de los procedimientos de extracción, control de calidad, limpieza y ensamble.

Palabras Clave: Microbioma Secuencias Ensamble Grafos

Contexto

El cuerpo humano es colonizado por una comunidad de microorganismos que se denomina microbioma y contiene diez veces más células que las suyas propias. La cantidad de genes presentes en total es varios órdenes de magnitud mayor que la del genoma humano. La nueva generación de tecnologías de secuenciación de ADN ha permitido estudiar las características del microbioma humano. El objetivo de estos estudios metagenómicos es analizar la estructura y la dinámica de las comunidades, para establecer cómo se relacionan sus miembros entre sí, cuáles son las sustancias que producen y consumen, y cómo se modifica la comunidad en presencia de enfermedades. El estudio por medio de la asignación funcional de cada gen del microbioma y su ubicación dentro del complejo de actividades metabólicas que ocurren en el paciente hospedador, en la comunidad microbiana, y en la interacción entre ambos, busca reconocer actividades metabólicas en el paciente asociadas con la presencia de enfermedades. Este es un campo de investigación muy activo con proyectos como el Metagenomics of the Human Intestinal Tract (MetaHIT), y abarca desde los aspectos médicos hasta el desarrollo y aplicación de nuevos algoritmos de explotación de datos y

reconocimiento de patrones. El objetivo general es entender el funcionamiento del microbioma humano a partir del procesamiento y análisis de muestras de secuencias de ADN, y construir nuevas herramientas de software para caracterizar el curso de patologías.

Introducción

El trabajo consiste en obtener los datos secuenciados de una muestra compuesta por varios microbiomas. El conjunto de secuencias ya ensambladas del microbioma habrá de compararse con otra base de datos de funciones genéticas para agrupar los genes integrantes por función y así obtener la distribución de frecuencias según las funciones metabólicas que las secuencias integrantes revelan [1]. Una vez formadas las matrices que representan por fila las distribuciones de cada microbioma individual estos pueden agruparse en clusters. Cada fila del conjunto representa a un paciente y en la base de datos esa instancia es un vector donde cada componente corresponde al número de genes del microbioma identificados con una dada función metabólica. Las características de cada agrupamiento logrado deben cotejarse con las apreciaciones clínicas de los pacientes que lo integran, ya obtenidas por otras vías diagnósticas, para apreciar el punto hasta el cual resultan útiles en la evaluación médica de la patología investigada.

Líneas de Investigación, Desarrollo e Innovación

En esta línea de trabajo ya se han estudiado procedimientos de obtención de datos y clustering de lo que se ha dado cuenta en presentaciones anteriores. En esta oportunidad se analizó el trabajo completo que se efectúa sobre las secuencias desde que son obtenidas por el secuenciador hasta que se construyen los

“contigs” mediante el ensamblado de secuencias que se realizó aplicando grafos. Los datos utilizados son nuevos pues se buscó que sirvieran luego para la identificación funcional. Su origen es la tecnología de secuenciación Illumina, y la forma de ensamblado es “de novo”, es decir sin que se utilicen genomas preexistentes y anotados como guía para ensamblar. Esta parte del trabajo se incluye dentro de los objetivos más generales que se intentan alcanzar en esta línea de investigación, desarrollo e innovación:

- Dominar la tecnología de almacenamiento, comparación y distribución funcional según las secuencias obtenidas del microbioma intestinal

- Determinar los métodos computacionales convenientes para los agrupamientos de microbiomas que revelen sus características clínicas.

- Establecer algoritmos de predicción entrenados y testeados para la evaluación clínica.

- Establecer una “pipeline” para la aplicación a pacientes locales

- Obtener muestras propias, enviarlas a secuenciar y aplicar los procedimientos probados.

Resultados y Objetivos

Los datos de secuenciación de este estudio son públicos y se busca poner a punto y automatizar, las operaciones bioinformáticas necesarias: selección de software más adecuado para cada una de dichas operaciones, análisis de calidad de las secuencias, diseño de una metodología de limpieza de las secuencias, validación, ensamblado de los metagenomas e integrar todos estos pasos.

El estudio cuenta con 143 muestras que se distribuyen de la siguiente manera: 16 de endoscopias, 99 de materia fecal y 28 de hisopados rectales. La secuenciación de

ADN total se realizó con tecnología Illumina con una estrategia de “paired-ends” de 300 nucleótidos, por lo que cada muestra está compuesta por dos archivos de secuencias. Esto significa que para cada fragmento de ADN analizado, se secuencian 300 nucleótidos desde cada extremo.

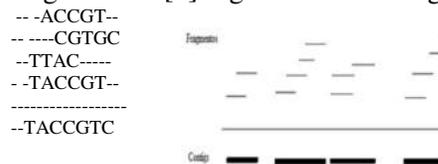
Todos los pasos se realizaron corriendo el sistema operativo Linux, distribución Ubuntu 16.04. Se procedió a la descarga de los 286 archivos utilizando la herramienta SRATOOLKIT de NCBI y se eliminaron dos muestras (4 archivos) que contaban con muy pocas secuencias. Las muestras restantes tenían entre, aproximadamente, 41600 y 521000 bases. Se realizó el control de calidad con el software FastQC y se determinó que casi todas las secuencias tenían restos de dos de los adaptadores que usa Illumina para la secuenciación, uno en la secuencia F (“forward”) y otro en la secuencia R (“reverse”) de cada “paired end”. Además se determinó que las frecuencias de cada base en los primeros 15 nucleótidos de las secuencias presentaban un nivel de variabilidad muy alto, que no era compatible con lo que se observaba más adelante y debido, posiblemente, a algún artefacto de la secuenciación que generaba “ruido”. También la calidad promedio de las secuencias caía por debajo del valor umbral que se fijó en 25 a partir de una posición que variaba para cada secuencia, pero que en general se ubicaba después de la posición 240. Para determinar el tipo exacto de secuencia contaminante y para tener una información más precisa del lugar donde ocurría se utilizó el software SCYTHE [3]. El proceso de limpieza se realizó con el programa CUTADAPT que efectúa la limpieza de adaptadores, cortes por caída en los valores de calidad, eliminación por largo mínimos, cortes en posiciones arbitrarias, etc. En primer lugar se

realizaron pruebas preliminares para determinar las opciones específicas de limpieza y los valores óptimos de los parámetros del programa. El proceso definitivo se efectuó en dos pasos. En el primero se eliminaron las secuencias contaminantes, se eliminó la parte 3’ de las secuencias que presentaran una caída en su calidad por debajo del valor umbral 25 mencionado antes y si alguna de las secuencias de un par “paired-end” después de estos cortes resultaba con una longitud menor a 50 bases se procedía a eliminar el par completo. En el segundo paso se eliminaron los primeros 15 nucleótidos del extremo 5’ y se volvieron a filtrar los pares para eliminar a aquellos con al menos un miembro de longitud menor a 50 bases.

Después de la limpieza se volvió a revisar la calidad de las secuencias con FastQC, con resultados satisfactorios.

Con las secuencias limpias y filtradas se procedió al paso de ensamblado. En la secuenciación el genoma es fragmentado. La longitud de cada fragmento depende de la tecnología empleada (≈ 600 bp con Illumina en paired end) y se necesita ensamblar los fragmentos luego de la secuenciación para identificar los genes de un individuo. La longitud en pares de bases (nucleótidos) de cada genoma varía según el tipo de organismo. Se entiende por read a cada fragmento obtenido del secuenciador. Varios reads se ensamblan para formar un contig que expresa los alineamientos de los nucleótidos que se encuentran en ellas.

La intención es que los contigs reconstruyan una parte de cada gen presente en la cadena de ADN que fue fragmentada [5] según se ve en la figura



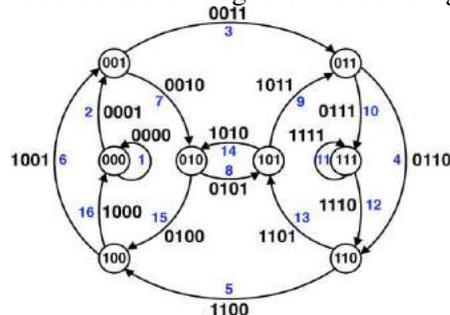
En 1735 L. Euler resolvió el llamado problema de los siete puentes [5]. Partiendo de una cualquiera de cuatro regiones, conectadas por siete puentes, en que quedaba dividida la ciudad, determinó las condiciones en las que cualquier problema similar puede resolverse afirmativamente. Euler demostró que:

- i) Si hay más de dos regiones a las cuales lleva un número impar de puentes entonces el camino buscado no existe
- ii) Si hay solo dos regiones unidas por un número impar de puentes el camino puede hacerse iniciándolo en cualquiera de ellas.
- iii) Si no hay ninguna región a la que lleva un número impar de puentes el camino siempre podrá hacerse iniciándolo en cualquier región. Modernamente se diría que para tratar el problema Euler representó cada puente con un arco del grafo y cada región con un nodo. Un ciclo euleriano comienza y termina en un mismo nodo pasando una sola vez por cada arco

En 1946 Nicolaas de Bruijn buscó resolver el problema llamado de la “supercadena”: encontrar la supercadena de caracteres más corta que contuviera a todas las posibles subcadenas de k símbolos (k -mers) de un alfabeto de n símbolos.

En un alfabeto de n caracteres existen subcadenas de k símbolos. Si los símbolos del alfabeto fueran las letras de los nucleótidos del ADN habría $4^3=64$ subcadenas posibles de tres nucleótidos (3-mers). Si en cambio el alfabeto estuviera formado por los símbolos binarios 0 y 1 todos los posibles 3-mers serían 000 001 010 011 100 101 110 y 111. Así se ve que la supercadena 0001110100 contiene a todos estos 3-mers. Es decir con 9 símbolos se

condensaron 24 según se ve en la figura.



De Bruijn imaginó esta solución por medio de grafos teniendo en cuenta los resultados de Euler [7] Si, por ejemplo, el alfabeto está formado por los símbolos 0 y 1 ($n=2$) y se quiere hallar la supercadena circular más corta que contenga a todos los k -mers con $k=4$, basta considerar los $(k-1)$ -mers ($k-1=3$) como los nodos de un grafo cuyos arcos dirigidos se constituyen tomando el prefijo del nodo de partida y el sufijo del nodo de llegada. Los distintos arcos así formados constituyen un ciclo euleriano. La supercadena cíclica resultante está formada por cada uno de los prefijos de los arcos. Es decir: 0000110010111101 A cada nodo salen y llegan en suma un número par de arcos con lo que se está en la condición iii) del teorema de Euler y por lo tanto el ciclo euleriano existe y es único.

En principio aplicar grafos al ensamblado de secuencias implicaría representar cada read por un nodo y los solapamientos entre lecturas por arcos.

Las técnicas de ensamblado de ADN utilizan un valor de $k \approx 55$ para fragmentos obtenidos por tecnología Illumina. Se pueden organizar las superposiciones de estas subsecuencias haciendo coincidir el subfijo de una inicial con el prefijo de la otra final. Esto se expresa en un grafo cuyos nodos son los k -mers y los arcos sus superposiciones. El ciclo hallado es Hamiltoniano pues pasa solo una vez por cada nodo.

$10^6=1000000$ es un número de lecturas que podría generar la secuenciación Illumina. Ellas requerirían unos 10^{12} alineamientos de k-mers y si hubiera 10^9 lecturas serían necesarios 10^{18} alineamientos. No existe algoritmo eficiente para esto pues la cantidad de operaciones a realizar no podría efectuarse en tiempos polinómicos. Como hallar un ciclo hamiltoniano es un problema NP-completo el proceso de ensamblado de ADN enfrenta a la computación con sus límites teóricos actuales.

En este punto entran Euler y de Bruijn. Resulta más sencillo y resoluble computacionalmente encontrar un ciclo que pase una sola vez por cada arco. Euler probó que si el grafo es conexo y no dirigido contiene exactamente un ciclo (euleriano) cuando cada nodo del grafo se asocia a un número par de arcos que lo conectan con otros nodos. En un grafo dirigido la cuestión es análoga: el número de arcos que parten del nodo tiene que ser igual al número de arcos que llega a él. (Grafo balanceado). En particular los grafos de de Bruijn contienen entonces un ciclo euleriano. Se trata de hallarlo.

No todos los supuestos de de Bruijn se cumplen en el caso de la fragmentación del genoma. En primer término no necesariamente se presentan todos los k-mers que podrían formarse. La solución que se ensaya es tomar $k \approx 55$ con la esperanza de que todos los arcos posibles estén efectivamente en el grafo. En segundo lugar algunos k-mers pueden repetirse frecuentemente. Para resolverlo se establece la "multiplicidad" del k-mer. Si su multiplicidad es m, se conecta su prefijo con su sufijo m veces y como el grafo también resulta balanceado existe el ciclo euleriano. Además, en general, la tecnología Illumina puede producir errores de lectura que, en tal caso, se intentan corregir antes del ensamblado.

Existe software libre que puede usarse para realizar el ensamblado de secuencias por esta vía. Por ejemplo el IDBA.UD desarrollado por el departamento de ciencias de la computación de la Universidad de Hong Kong [8] y [9]

Formación de Recursos Humanos

En el equipo de trabajo participan un magister en explotación de datos y otra en bioinformática, un doctor en biología, una médica, 2 ingenieros en sistemas, una matemática y un estudiante de ingeniería informática. Está en curso una tesis de doctorado y otra de maestría.

Referencias

- [1] Arumugam, M et al. Enterotypes of the human gut microbiome. Nature 2011 may 12; 473(7346): 174-180. doi:10.1038/nature09944
- [2]<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- [3] <https://github.com/vsbuffalo/scythe>
- [4](<https://cutadapt.readthedocs.io/en/stable/>)
- [5] Lander ES, Waterman MS (1988). "Genomic mapping by fingerprinting random clones: a mathematical analysis". Genomics. 2 (3): 231-239. doi:10.1016/0888-7543(88)90007-9. PMID 3294162.
- [6] Euler, L. Commentarii Academiae Scientiarum Petropolitanae 8, 128-140 (1741)
- [7] Campea, Ph E C; Pevzner, P A; Tesler, G. How to apply de Bruijn graphs to genome assembly. Nature Biotechnology Volume 29 Number 11. 987-991. (2011)
- [8] Peng, Y; Leung, H C M; Yiu, S M; Chin F Y L. IDBA-UD: de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics. Vol. 28. n° 11. 1420-1428. (2012)
- [9] i.cs.hku.hk/~alse/hkubrg/projects/idba_ud/

Calidad Universitaria mediante Técnicas del Data Mining

Myriam Herrera¹, María Inés Lund², Susana Beatriz Ruiz¹, María Gema Romagnano²

¹Departamento de Informática, Facultad de Ciencias Exactas, Físicas y Naturales,
Universidad Nacional de San Juan

²Instituto de Informática, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad
Nacional de San Juan

mherrera, mlund, mromagnano@iinfo.unsj.edu.ar, sbruizr@yahoo.com.ar

RESUMEN

En la actualidad, la mayoría de las instituciones, empresas u organizaciones miden la calidad de sus productos y/o servicios. De igual forma las instituciones educativas se ven obligadas a medir la calidad educativa. Para ello se necesita conocer los factores que influyen en la calidad de la institución, entre ellos, los relacionados al rendimiento académico de sus alumnos y al grado de satisfacción de sus egresados. Para lograr esto se utilizarán valiosas técnicas estadísticas que permitirán clasificar sujetos u objetos a partir de características similares. Estas técnicas se pueden diferenciar por la manera de extraer conocimiento útil escondido en los datos. Por un lado, el Análisis Discriminante, también referido como reconocimiento de patrones supervisado o asistido o aprendizaje con guía. Por otro lado, el Análisis de Conglomerados, referido como reconocimiento de patrón no supervisado o conocimiento sin guía. Como es común recopilar grandes conjuntos de datos, de distinta naturaleza, en voluminosas bases de datos, es que se utilizarán los análisis de Datos Simbólicos empleando la Lógica Difusa, que son también herramientas para Data Mining. En este proyecto se aplicarán las técnicas mencionadas para analizar los factores influyentes en la calidad universitaria como así también se detectarán tipologías básicas de grupos, obtenidos de los alumnos universitarios y egresados de la Facultad de Ciencias Exactas, Físicas y Naturales de la UNSJ.

Palabras clave: Calidad Universitaria, Clasificación, Data Mining.

CONTEXTO

Este proyecto ha sido presentado en la Convocatoria 2017 de la Universidad Nacional de San Juan, para el periodo comprendido entre 01/01/2018 al 31/12/2019 y se encuentra en proceso de evaluación externa.

Se desarrollará en el ámbito de la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de San Juan, con el apoyo de la Secretaría de Asuntos Estudiantiles, el Instituto de Informática y los Departamentos de Biología, de Geofísica y Astronomía, de Informática y de Geología, que abarcan las distintas carreras que se dictan en esta Facultad.

Es un proyecto que continúa en línea con las investigaciones desarrolladas en proyectos anteriores, entre los que se mencionan:

- “Técnicas de Clasificación aplicadas al rendimiento académico”. Acreditado por el CICITCA. Vigencia: 01/01/2016 – 31/12/2017. Código: 21 E / 1011.
- “Algoritmos de Clasificación de Procesos Multivariados utilizando Medidas de Asociación Espacial”. Acreditado por el CICITCA. Vigencia: 01/01/2014 – 31/12/2015. Código: 21 E / 948.
- “Determinación y Comparación de Perfiles Sociales y Culturales de Estudiantes Universitarios a través de Técnicas Estadísticas Multivariadas”. Acreditado por el CICITCA. Vigencia:

- 01/01/2014 – 31/12/2015. Código 21 F/ 982.
- “Clasificación Espacial Multivariada”. Acreditado por el CICITCA. Vigencia: 01/01/2011 – 31/12/2013. Código: 21 E/ 878
 - “Reducción y Selección de Variables en la Clasificación Digital”. Acreditado por el CICITCA. Vigencia: 1/01/2008 - 31/12/2010. Código: 21 E/ 820.
 - “Aplicación de una metodología en la medición de la calidad del proceso enseñanza aprendizaje en la universidad”. Acreditado por el CICITCA. Vigencia: 01/05/2003 al 31/12/2005.

1. OBJETIVOS

Objetivo General

Determinar factores influyentes en los alumnos y egresados universitarios que caractericen la Calidad Universitaria.

Objetivos Específicos

- Formar recursos humanos, al nivel de grado y posgrado, en la temática que involucra la Gestión de Calidad en Educación y específicamente en la metodología elaborada.
- Analizar los datos otorgados por los sistemas SIU Kolla (Egresados) y SIU Guaraní (Alumnos).
- Determinar las variables influyentes que caractericen a alumnos y egresados universitarios.
- Identificar qué variables tienen mayor poder de discriminación y de predicción en la clasificación de sujetos.
- Determinar los datos simbólicos que determinarán reglas lógicas y taxonomías de las variables influyentes en alumnos y egresados.
- Transferir la herramienta metodológica obtenida a instituciones educativas del medio.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Los datos se han convertido en un recurso crítico en muchas organizaciones e instituciones con diversos objetivos y por lo tanto, el acceso eficiente a estos, el

compartirlos, extraer información de los mismos y hacer uso de la información se transforma en una urgente necesidad. Existen varios enfoques de investigación que han aportado en ésta temática [1], [2], [3], [4]. El objetivo de analizar y comprender grandes y complejos conjuntos de datos, que posteriormente conducen a valiosa información, es común a todos los campos de los negocios, ciencia, ingeniería, entre otros [5]. La habilidad para extraer conocimiento útil escondido en esos datos y actuar sobre ése conocimiento está transformándose en algo cada vez más importante en el mundo competitivo actual. Como resultado hay muchos esfuerzos, no sólo para integrar varias fuentes de datos dispersos a través de sitios diferentes, sino también extraer información de esas bases de datos en la forma de patrones y tendencias. Dentro de la Inteligencia Artificial, la Minería de Datos, comúnmente conocida como Data Mining, analiza conjuntos de datos para encontrar relaciones y resúmenes de datos útiles para el propietario de los datos. Estas relaciones y resúmenes derivados a través del ejercicio del Data Mining se refieren a modelos y patrones.

La aplicación automatizada de algoritmos de minería de datos permite detectar fácilmente patrones en los datos. Los algoritmos de minería de datos se clasifican en dos grandes categorías: supervisados o predictivos y no supervisados o de descubrimiento del conocimiento [6], [7].

Los modelos supervisados o predictivos requieren de un conjunto de pruebas y de interacciones de entrenamiento. Las técnicas usadas son la clasificación (Análisis Discriminante) y la predicción de valores. Los modelos no supervisados o descriptivos descubren patrones y tendencias en los datos actuales (no utilizan datos históricos). Las técnicas usadas son: Asociación, Segmentación o 'Clustering' (Análisis de Conglomerados) [8], [9].

El Reconocimiento de Patrones tiene como objetivo la **clasificación** de objetos dentro de un número de categorías o clases. Dependiendo de la aplicación estos objetos pueden ser imágenes, señales o cualquier tipo

de medidas que necesitan ser clasificadas. Esas medidas se llaman patrones.

Las medidas usadas para la clasificación de objetos o patrones son conocidas como características. El conjunto de todas las características forman el vector que identifica únicamente a un patrón (objeto).

Las cuestiones que preocupan en el diseño de un sistema de clasificación que ejecuta la tarea de un reconocimiento de patrones dados son: (a) la generación de características para lo cual es importante la elección del mejor número de características; (b) el diseño del clasificador y finalmente, cuando el clasificador está diseñado, (c) la evaluación del rendimiento del clasificador diseñado mediante el error de clasificación [10], [11], [12].

Análisis de Datos Simbólicos es también una herramienta para **Data Mining** que generaliza los métodos clásicos exploratorios e informáticos. En muchas actividades humanas es común recopilar considerables conjuntos de datos en grandes bases de datos, por lo cual es importante resumir estos datos en términos de sus conceptos con el sentido de extraer nuevos conocimientos. Estos conceptos se pueden describir por tipos de datos más complejos, llamados Datos Simbólicos, que contienen variación interna y son estructurados. Es en este contexto que surge la necesidad de extender los métodos de análisis de datos estándar (exploratorio, representaciones gráficas, clustering, análisis factorial, discriminación, etc.) a estos datos simbólicos. Los datos simbólicos implican tablas de datos más complejas llamadas tablas de datos simbólicos. Una celda de tales tablas no necesariamente contiene valores categóricos o cuantitativos simples, sino muchos valores, que pueden tener pesos o estar unidos por reglas lógicas y taxonomías. Por ejemplo, una celda puede contener un intervalo o una distribución. Este tipo de Análisis será generado para analizar y tomar decisiones sobre grandes bases de datos, especialmente para datos de encuestas. Si bien resumen, en gran proporción, las bases, éstas preservan lo esencial o la información de

interés. Además, permiten visualizar, comparar y clasificar objetos [10], [13], [14]. También el uso de la Lógica Difusa puede ser de vital importancia en cualquier proceso de Minería de Datos ya que es habitual que el conjunto de datos a analizar se haya obtenido con un propósito distinto al de la extracción de conocimiento. Es común la presencia de información numérica junto con información textual, con ambigüedades por el uso de diferentes símbolos con igual significado, redundancia, términos perdidos, imprecisos o erróneos, etc.[15]. La inclusión de la Lógica Difusa dentro del Soft Computing, constituye una herramienta de representación del conocimiento que permite modelar incertidumbre e imprecisión de una forma sencilla y directamente interpretable por el usuario [16]. En este proyecto se aplicarán las técnicas mencionadas, o una combinación de ellas.

3. FORMACION DE RECURSOS HUMANOS

En el grupo de trabajo se encuentran dos doctorandos. Se prevé la incorporación al presente proyecto de profesionales que actualmente están cursando posgrados en Computación, y también no docentes y alumnos ayudantes.

4. REFERENCIAS

- [1] M. de M. Diaz, P. A. Urquijo, J. M. Arias Blanco, T. Escudero Escorza, S. Rodriguez Espinar, and J. Vidal García, "Evaluación del rendimiento en la enseñanza superior. Comparación de resultados entre alumnos procedentes de la LOGSE y del COU," in *Revista de Investigación Educativa*, vol. 20, no. 2, M. de Miguel, P. A. Urquijo, J. M. A. Blanco, T. E. Escorza, S. R. Espinar, and J. V. García, Eds. 2002, pp. 357–383.
- [2] T. Escudero Escorza, "La evaluación y mejora de la enseñanza en la Universidad: otra perspectiva," in *Revista de Investigación Educativa*, vol. 18, no. 2, 2000, pp. 405–416.
- [3] G. M. Garbanzo Vargas, "Factores asociados al rendimiento académico en

- estudiantes universitarios, una reflexión desde la calidad de la educación superior pública,” *Rev. Educ.*, vol. 31, no. 1, pp. 43–63, 2007.
- [4] M. H. Efrón and A. Pérez Lindo, Eds., *Aportes al debate sobre la gestión universitaria I*. Buenos Aires, Argentina: Editorial De los cuatro vientos, 2005.
- [5] N. Moscoloni, “Las Nubes de Datos. Métodos para analizar la complejidad,” 2005.
- [6] A. . Anaya and J. G. Boticario, “A data mining approach to reveal representative collaboration indicators in open collaboration frameworks,” 2009.
- [7] M. de L. Gurmendi, “De la información al conocimiento: factores que ayudan a un mejor uso de la tecnología en la gestión - Nülan,” in *Aportes al debate sobre la gestión universitaria II*, R. I. In Efrón, Marcelo Héctor y Vega, Ed. Buenos Aires: De los cuatro vientos, 2005, pp. 66–75.
- [8] I. González López, “Realización de un Análisis discriminante explicativo del rendimiento académico en la Universidad,” in *Revista de Investigación Educativa*, vol. 22, no. 1, 2004, pp. 43–59.
- [9] M. M. Torrado-Fonseca and V. Berlanga-Silvente, “Revista d’innovació i recerca en educació,” *Rev. d’Innovació i Recer. en Educ.*, vol. 6, no. 2, pp. 150–166, 2013.
- [10] H.-H. Bock and E. Diday, *Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information from Complex Data*. Springer Berlin Heidelberg, 2000.
- [11] D. Peña, *Análisis de datos multivariantes*. McGraw-Hill/Interamericana, 2002.
- [12] “Spad - Software Informer. SPAD,a data analytics software, uses company data to anticipate risks.” [Online]. Available: <http://spad.software.informer.com/>. [Accessed: 13-Mar-2017].
- [13] E. Diday, “Análisis de Datos Simbólicos,” *Rev. IRICE*, vol. 11, 1997.
- [14] R. A. (Richard A. Johnson and D. W. Wichern, *Applied multivariate statistical analysis*. Pearson Prentice Hall, 2007.
- [15] J. Hernández Orallo, M. J. Ramírez Quintana, and C. Ferri Ramírez, *Introducción a la Minería de Datos*. Editorial Pearson, 2004.
- [16] L. A. Zadeh, “Fuzzy Sets,” *Inf. Control*, vol. 8, pp. 338–353, 1965.

AVANCES EN SELECCIÓN DE BIOMATERIALES UTILIZADOS EN IMPLANTES DENTALES APLICANDO TÉCNICAS DE MINERÍA DE DATOS

Ganz, Nancy B.^(a); Domínguez, Facundo A.^(b); Ares Alicia E.^(a); Kuna, Horacio D.^(b)

(a) Programa de Materiales y Físicoquímica; Laboratorio de Ciencia de los Materiales; Instituto de Materiales de Misiones; CONICET; Facultad de Ciencias Exactas Químicas y Naturales; Universidad Nacional de Misiones.

(b) Departamento de Informática; Facultad de Ciencias Exactas Químicas y Naturales; Universidad Nacional de Misiones.

nancy.bea.ganz@gmail.com

RESUMEN

En Argentina, como en otros países del mundo se fabrican biomateriales para diferentes aplicaciones con el objetivo de restaurar las funciones del cuerpo humano, como es el caso de los implantes dentales. Sin embargo, es necesario contar con información suficientemente calificada y accesible sobre: tipos de implantes dentales, propiedades de los biomateriales, causas de fallo, características y condiciones de salud de los pacientes que requieren de estos implantes. La carencia de un registro digital con datos reales de historias clínicas de pacientes que se han sometido a procesos quirúrgicos de colocación de implantes dentales, dificulta la tarea de análisis y extracción de conocimiento oculto para los especialistas implantólogos, sobre la relación implante dental - condición del paciente. De aquí, surge la necesidad de crear un registro automatizado que reúna las variables que representan el proceso de colocación de un implante dental, con el objetivo de identificar los factores que contribuyen al éxito o al fracaso de los implantes dentales colocados en la Provincia de Misiones, a través de la aplicación de técnicas de minería de datos, mediante el diseño de un procedimiento con métodos híbridos, y la aplicación de la metodología CRISP-DM para asentar el proceso.

Palabras Clave: *Minería de Datos, Biomateriales, Implantes Dentales.*

CONTEXTO

Esta línea de investigación se lleva a cabo dentro del Programa de Materiales y Físicoquímica (PROMyF) en el Laboratorio de Ciencia de los Materiales del Instituto de Materiales de Misiones (IMAM), de la Facultad en Ciencias Exactas, Químicas y Naturales (FCEQyN), de la Universidad Nacional de Misiones (UNaM), en el marco de un plan de tesis doctoral. Está financiado por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) a través de una “Beca Interna Doctoral” otorgada por Resolución D N° 4869. Además, cuenta con el auspicio del Colegio de Odontólogos de la Provincia de Misiones.

1. INTRODUCCIÓN

La falta de un registro digital de implantes dentales con datos sobre condiciones del paciente, el proceso y el implante, hace dificultosa la tarea de investigación para extraer conocimiento que podría llegar a ayudar al proceso de oseointegración de un implante.

Los implantes dentales son biomateriales, es decir son materiales biológicos que reemplazan y/o restauran la función de un tejido o de un órgano vivo (pieza dental). Son inertes, debido a que están diseñados para ser implantado tanto dentro como fuera del cuerpo humano. [1]

El proceso de oseointegración define el éxito del implante, y éste depende de tres factores principales: propiedades y biocompatibilidad del implante, condiciones de salud del paciente, y la habilidad del cirujano. [2]

Sin embargo, la rehabilitación oral a través de implantes dentales puede presentar riesgos relacionados con la etapa del proceso de oseointegración implante/hueso. Estos riesgos pueden estar relacionados con las condiciones de salud del paciente, la técnica quirúrgica empleada por el especialista implantólogo, el tipo de implante, así como el tabaquismo. [3]

Actualmente, el campo de la minería de datos ha tenido muchos avances respecto a la aplicación y desarrollo de técnicas en el sector de la salud, para la predicción de enfermedades y para la toma de decisiones en base al análisis de grandes cantidades de datos. Por ejemplo: [4] Extrae datos de Historias Clínicas Electrónicas y utiliza la combinación de los algoritmos J48 y K-Means, para contribuir al diagnóstico de la hipertensión arterial. [5] Compara Árbol de Decisión, Clasificación Bayesiana, métodos predictivos como K-Nearest Neighbors (KNN) y Redes Neuronales para la predicción de enfermedad cardíaca. [6] Utiliza Redes Neuronales y Reglas de Asociación, para la detección y clasificación de tumores en mamografía digital. [7] Utiliza Redes Neuronales para discriminar grupos entre pacientes que se encontraban en terapia intensiva. [8] Utiliza Redes Bayesianas y Árboles de Decisión, además evalúa la utilidad de la metodología bayesiana en la predicción y el diagnóstico médico de enfermedades complejas. En [9] presenta la aplicación de métodos de aprendizaje automático (Red Neuronal, Máquina de Vector Soporte y Árbol de Decisión) al problema de predicción de caries en niños. En [10] presenta un estudio de minería de datos para formular nuevas hipótesis sobre la asociación entre la periodontitis y la diabetes tipo II.

Existe una gran cantidad de trabajos de minería de datos aplicados a la detección temprana y tratamiento de enfermedades oncológicas, así como en la identificación, prevención y

reducción de riesgos. Sin embargo, no se han encontrado trabajos que apliquen técnicas de minería de datos al campo de los biomateriales, específicamente a los implantes dentales, en [11] se muestra un análisis estadístico de Regresión Logística Múltiple para determinar los factores que influyen en el éxito de los implantes dentales. En [12] hace un estudio donde determina si existe relación entre los fracasos de los implantes dentales y las enfermedades sistémicas (concretamente sobre la osteoporosis, hipertensión, diabetes e hipotiroidismo), en un población de pacientes sometidos a cirugía de implantes dentales en el hospital San José de Santiago de Chile. En [13] presenta un estudio comparativo de tres técnicas de aprendizaje automático (Máquina de Vector Soporte (SVM), Máquina de Vector Soporte Ponderado y Redes Neuronales RBF constructivas (RBF-DDA) con parámetro de selección) para la predicción del éxito de los implantes dentales.

Si bien estos trabajos utilizan datasets de implantes dentales, no centran su atención al biomaterial, es por esto que esta línea de investigación busca abordar conjuntamente el estudio de las características del implante utilizado.

La detección de datos anómalos usando técnicas de Minería de Datos permite detectar grupos de datos que pueden ser de especial interés en temas relacionados como por ejemplo con el análisis de los biomateriales.[14]

Para documentar proyectos de Minería de Datos, existen procesos diseñados para guiar al usuario a través de una sucesión formal de pasos. Entre las metodologías más importantes encontramos: SEMMA, CRISP-DM [15] y P3TQ.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Es de vital importancia un registro digital con datos de historias clínicas sobre la colocación de implantes dentales, con variables que representen información relacionada al implante y a las condiciones de salud del paciente. En este contexto se reconoce como objetivo de interés para la presente línea de investigación, la creación de un registro digital de implantes dentales y el desarrollo de un procedimiento a través de una metodología híbrida, para facilitar la extracción de conocimiento útil para el especialista y actores de interés, e identificar los factores que contribuyan al éxito o al fracaso de los implantes dentales, así como las condiciones óptimas que debe tener el paciente y el implante dental.

Además, estudiar las propiedades mecánicas, químicas y físicas de los biomateriales utilizados en la implantología dental, para una mejor comprensión de la funcionalidad y resistencia de los implantes. Conjuntamente, lograr una taxonomía según el origen de fabricación y según el tipo de biomaterial utilizado en la industria del implante.

3. RESULTADOS OBTENIDOS/ESPERADOS

Entre los resultados obtenidos, destacamos la creación de un registro digital (datasets) con datos reales de historias clínicas de pacientes que se han sometido al proceso quirúrgico de colocación de implantes dentales en las localidades más características de la Provincia de Misiones.

Así mismo, se está trabajando en conjunto con expertos en el área de Minería de Datos, Biomateriales, Especialista en Implantología y Rehabilitación Compleja, Especialista en Patología Bucal e Implantología Oral, así como en Estadística.

Para comenzar con el proceso de extracción de conocimiento, los datos utilizados en este trabajo fueron obtenidos manualmente a través del diseño y corroboración de un formulario por parte de los expertos en el área. Así como, sucesivas entrevistas y charlas con más de 30

especialistas en implantología oral de toda la provincia de Misiones, con el fin de determinar cuáles son los datos de importancia a la hora de considerar la colocación de un implante dental.

Los datos del datasets se agrupan en 4 secciones:

- **Datos del Paciente:** referidos a los antecedentes y condiciones médicas de los pacientes a la hora de la intervención.
- **Datos del Implante:** referidos a las características del implante utilizado por el especialista implantólogo.
- **Datos de la Fase Quirúrgica:** referidos al procesamiento de intervención quirúrgica y mejoramiento del lecho óseo del paciente.
- **Datos del Seguimiento Postoperatorio:** referidos al resultado del proceso de colocación del implante, es decir si el proceso fue exitoso o fracaso.

Estas 4 secciones se encuentran alojadas en un solo archivo de datos, denominado BD_IMPLANTES.csv.

En base a la metodología CRISP-DM, la *fase 1 de comprensión del problema* abarco la definición del objetivo del presente trabajo, el cual es la determinación de los factores que contribuyen al éxito de los implantes dentales, así mismo se evaluó la situación de la carencia de un registro de estas características, se fijó el objetivo de minería de datos que es el diseño de un procedimiento a través de una metodología híbrida. Se precisó el plan de trabajo, que consistió en el relevamiento bibliográfico, investigación de la situación actual de la implantología en la provincia, definición de los expertos en el área, iniciar la recolección de datos y comenzar con la preparación de estos datos. En la *fase 2 de comprensión de los datos* se plasmó la recolección de los datos iniciales, donde se observa que los datos se agrupan en 4 secciones: datos del paciente, datos del implante, datos de la fase quirúrgica y datos del seguimiento postoperatorio. Así mismo, el

datasets cuenta con un total de 34 variables y 1050 casos de implantes colocados en la provincia de Misiones. Existen variables cuantitativas (como por ejemplo: *edad*, *longitud* y *diametro*) y variables cualitativas categóricas (como por ejemplo: *nivel_prof*, *pieza_dentaria*, *protoc_carga*, *seguir_postop*, entre otras). Seguidamente se detalla el contenido (columnas) del datasets “*BD_IMPLANTES.csv*” correspondiente a cada sección:

<i>Datos del Paciente</i>	edad, genero, ocupacion, o_s, enfermedad, fuma, alcoholismo, periodontitis, riesgo_periodontal, ingest_med, alergia.
<i>Datos del Implante</i>	material, marca, diseño, longitud, diametro, conexion, procedencia.
<i>Datos de la Fase Quirúrgica</i>	fecha_inter, lugar_proc, nivel_prof, pieza_dentaria, protoc_carga, exodoncia, expan_osea, elev_seno_maxilar, reg_tejidos_duros, reg_tejidos_blandos, tiempo_coloc, tipo_hueso,

	indic_protésica, comp_quirúrgica.
<i>Datos de Seguimiento Postoperatorio</i>	seguir_postop, observacion.

Actualmente, esta línea de investigación se encuentra en la etapa de preparación de los datos y verificación de la calidad de los mismos, mediante la aplicación de filtros.

Además, se busca la transformación y reducción de las variables que no aportaban ganancia de información al estudio de caso, con el fin de construir un datasets final para iniciar con el diseño de un procedimiento que utilice técnicas de minería de datos para resolver el problema planteado, e ir ajustando en función de las necesidades que se presenten.

Conjuntamente, se está profundizando el estudio y adaptabilidad del algoritmo máquina de vector soporte (SVM), esto en función de los trabajos citados que utilizan y recomiendan la aplicación de este método para datasets de este tipo.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto es parte de las líneas de investigación del “Programa de Materiales y Físicoquímica” de la FCEQyN – UNaM, con cuatro integrantes; de los cuales dos son docentes de la FCEQyN – UNaM (donde uno pertenece al programa mencionado y es investigador independiente del CONICET y el otro es investigador categoría I perteneciente al Departamento de Informática). El tercer integrante se encuentra realizando un doctorado; el cuarto y último integrante se encuentra desarrollando su tesina de grado de la carrera de Licenciatura en Sistemas de Información de la FCEQyN – UNaM.

5. BIBLIOGRAFÍA

- [1] B. D. Ratner, A. S. Hoffman, F. J. Schoen, and J. E. Lemons, *Biomaterials Science. An introduction to Materials in*

- Medicine. San Diego, California - USA: Elsevier Academic Press, 1996.
- [2] G. P. Kothiyal and A. Srinivasan, "Trends in Biomaterials", *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2016.
- [3] P. Laine, A. Salo, R. Kontio, S. Ylijoki, C. Lindqvist, and R. Suuronen, "Failed dental implants – clinical, radiological and bacteriological findings in 17 patients", *J. Cranio-Maxillofacial Surg.*, vol. 33, no. 3, pp. 212–217, Jun. 2005.
- [4] F. Dávila Hernández and Y. Sánchez Corales, "Técnicas de minería de datos aplicadas al diagnóstico de entidades clínicas", *Rev. Cuba. Informática Médica*, vol. 4, no. 2, pp. 174–183, 2012.
- [5] J. Soni, U. Ansari, D. Sharma, and S. Soni, "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction", *Int. J. Comput. Appl.*, vol. 17, no. 8, pp. 43–48, 2011.
- [6] M. Antonie, A. Coman, and O. R. Zaiane, "Application of Data Mining Techniques for Medical Image Classification", *Proc. Second Int. Work. Multimida Data Min.*, pp. 94–101, 2001.
- [7] C. Arizmendi, E. Romero, R. Alquezar, P. Caminal, I. Díaz, S. Benito, and B. Giraldo, "Data mining of patients on weaning trials from mechanical ventilation using cluster analysis and neural networks", *2009 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 4343–4346, 2009.
- [8] G. Solarte and Y. Castro, "Modelo híbrido para el diagnóstico de enfermedades cardiovasculares basado en inteligencia artificial", *Rev. Tecnura*, vol. 16, no. 33, pp. 35–52, 2012.
- [9] R. D. Montenegro, A. L. I. Oliveira, G. Cabral, C. R. T. Katz, and A. Rosenblatt, "A comparative study of machine learning techniques for caries prediction", *20th IEEE Int. Conf. Tools with Artif. Intell.*, vol. 2, pp. 477–481, 2008.
- [10] U. Covani, S. Marconcini, G. Derchi, A. Barone, and L. Giacomelli, "Relationship between human periodontitis and type 2 diabetes at a genomic level: a data-mining study", *J. Periodontol.*, vol. 80, no. 8, pp. 1265–1273, 2009.
- [11] J. E. B. Tamez, F. N. Zilli, L. A. Fandiño, and J. M. Guizar, "Factores relacionados con el éxito o el fracaso de los implantes dentales colocados en la especialidad de Prosthodontia e Implantología en la Universidad de La Salle Bajío", *Revi. Esp. Cirugía Oral y Maxilofac.*, pp. 1–9, 2016.
- [12] J. Domínguez, J. Acuña, M. Rojas, J. Bahamondes, and S. Matus, "Study of association between systemic diseases and dental implant failure", *Rev. Clínica Periodoncia, Implantol. y Rehabil. Oral*, vol. 6, no. 1, pp. 9–13, 2013.
- [13] A. L. I. Oliveira, C. Baldisserotto, and J. Baldisserotto, "A comparative study on machine learning techniques for prediction of success of dental implants", *Mex. Int. Conf. Artif. Intell.*, vol. 3789 LNAI, pp. 939–948, 2005.
- [14] H. Kuna, R. G. Martinez, and F. Villatoro, "Automatic Outliers Fields Detection in Databases", *J. Model. Simul. Syst. HyperSciences Publ.*, vol. 3, no. 1, p. 14–20., 2012.
- [15] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer and R. Wirth, *CRISP-DM 1.0: Step-by-step data mining guide*, 2000.

Estructuras de Datos y Algoritmos Eficientes para Búsquedas Web y Procesamiento de Grandes Datos

Gabriel H. Tolosa¹, Pablo Lavallén¹, Tomás Delvechio¹, Agustín Marrone¹,
Francisco Tonin Monzon¹, Federico Radeljak¹, Esteban A. Rissola^{1,3} y Esteban Feuerstein²
{tolosoft, plavallen, tdelvechio, eamarrone, ftonin, fradeljak}@unlu.edu.ar
esteban.andres.rissola@usi.ch; efeurest@dc.uba.ar

¹Departamento de Ciencias Básicas, Universidad Nacional de Luján

²Departamento de Computación, FCEyN, Universidad de Buenos Aires

³Facoltà di Scienze Informatiche, Università della Svizzera Italiana

Resumen

Desde hace ya un par de décadas, la cantidad de información, las aplicaciones y el número de usuarios digitales crece exponencialmente, formando un ecosistema en el cual se intenta explotar la masividad de los datos y presentan a la comunidad científico/tecnológica nuevos desafíos.

Por ejemplo, los motores de búsqueda para la web son aplicaciones que procesan miles de millones de documentos para responder consultas de los usuarios. Esto genera nuevas necesidades de almacenamiento, procesamiento y búsquedas, expandiendo los límites del trabajo en una sola máquina y unos pocos algoritmos. Las consultas deben responderse en milisegundos, con resultados relevantes, sobre un escenario altamente heterogéneo.

Además, el área de “Big Data”, que se aplica a cúmulos de datos que no pueden ser procesados y/o analizados de forma eficaz y eficiente utilizando técnicas tradicionales, aporta nuevos enfoques que complementan la idea anterior.

En este proyecto se estudian, proponen, diseñan y evalúan estructuras de datos y algoritmos para soportar búsquedas de escala web y/o analizar datos masivos de forma eficiente.

Palabras clave: algoritmos eficientes, motores de búsqueda, estructuras de datos, grandes datos.

Contexto

Esta presentación se encuentra enmarcada en los proyectos de investigación “Algoritmos Eficientes y Minería Web para Recuperación de Información a Gran Escala” del Departamento de Ciencias Bási-

cas (UNLu) y “Modelos y herramientas algorítmicas avanzadas para redes y datos masivos” del Departamento de Computación de la Facultad de Ciencias Exactas y Naturales (UBA).

Introducción

Desde hace ya un par de décadas, la información digital está creciendo exponencialmente. Algunas estimaciones indican que desde 2005 hacia 2020 el universo digital multiplicará por un factor de 300 su crecimiento, duplicando su tamaño aproximadamente cada 2 años [15]. A modo de ejemplo, en Twitter¹ se publican más de 250.000 tweets por segundo y en Facebook se generan más de 40.000 posts (también por segundo), que equivalen a más de 300 GB. El también creciente número de usuarios en el mundo digital, en particular, en redes sociales utilizando dispositivos móviles [10, 11], conforman un ecosistema en el cual se desarrollan nuevas actividades que intentan explotar la masividad de los datos y presentan a la comunidad científico/tecnológica nuevos desafíos.

Por esto, existe la necesidad permanente de nuevos enfoques, estrategias y técnicas que, haciendo uso de las herramientas computacionales adecuadas, ayuden a resolver de forma eficiente los problemas que aparecen continuamente. El ejemplo más saliente a mencionar es el caso de los motores de búsqueda para la web, herramientas que se basan en un necesidad de información del usuario e intentan satisfacerla. En esta área, se aplican técnicas de re-

¹<https://blog.hootsuite.com/twitter-statistics/>

cuperación² sobre una porción del espacio web que han recorrido y procesado [2].

El tamaño y complejidad de la web³ genera nuevas necesidades de almacenamiento, procesamiento y búsquedas, expandiendo los límites del trabajo en una sola máquina y unos pocos algoritmos. Se requiere hoy de procesamiento distribuido, paralelo y altamente eficiente. Las consultas deben ser respondidas en pequeñas fracciones de tiempo (milisegundos) y deben ofrecer resultados relevantes sobre un escenario altamente heterogéneo, en el cual aparecen oportunidades únicas para avances científico/tecnológicos en áreas como algoritmos, estructuras de datos, sistemas distribuidos y procesamiento de datos a gran escala [5].

Como parte también de esta realidad, uno de los conceptos que ha ganado espacio en los últimos años es el de “Big Data”. En términos generales, representa la idea de la masividad de datos, su velocidad de aparición y variedad de fuentes (estructurados, semi-estructurados y no estructurados) que se generan y requieren ser procesados para diferentes tareas. En particular, no son viables para el almacenamiento en sistemas de gestión de bases de datos tradicionales y, menos aún, en un único equipo de cómputo. Por lo tanto, esta idea aplica a cúmulos de datos que no pueden ser procesados y/o analizados de forma eficaz y eficiente utilizando técnicas y herramientas tradicionales [19], por lo que se requieren nuevos enfoques. En la actualidad, muchas aplicaciones importantes requieren de estrategias derivadas del concepto de Big Data para el procesamiento de sus datos. Por ejemplo, las redes sociales [20], los servicios de *streaming* [34], la genómica [24], la meteorología [16], entre otros.

Para el análisis de grandes volúmenes de datos, consumidos por millones de usuarios en casi todos los ámbitos de la actividad humana se emplean, entre otras, técnicas del área de *machine learning* como reconocimiento de patrones en textos, análisis estadístico, visualización, agrupamientos, redes neuronales, etc. [25], las cuales permiten convertir datos en información útil. En algunos casos, los problemas aumentan su complejidad ya que en sus procesos ingestan grandes volúmenes de datos de fuentes diversas en tiempo real [26].

²En sentido amplio, ya que se trata de múltiples fases de procesamiento como búsquedas y ranking, entre otras.

³Por ejemplo, Google indexa unos 45.000 millones de documentos de acuerdo a <http://www.worldwidewebsize.com>

Líneas de I+D

Este proyecto sigue con las líneas de I+D del grupo, las cuales se dividen en dos grupos principales, aunque con temas en la intersección de ambos. En el primer caso, se trata de mejorar la eficiencia en la recuperación de información de gran escala, rediseñando los algoritmos internos y las estructuras de datos usadas. En el segundo, se proponen y evalúan arquitecturas de procesamiento de grandes datos para problemas diversos. En ambos casos, existen oportunidades de investigación en temas poco explorados por la comunidad científica. En particular, las líneas de I+D principales son:

a. Estructuras de Datos y Algoritmos para Búsquedas

Los sistemas de búsqueda en texto utilizan como estructura de datos básica el índice invertido, formado por un vocabulario (V) con todos los posibles términos y un conjunto de *posting lists* (L) con información acerca de los documentos donde aparece cada término junto con información usada para el ranking (por ejemplo, la frecuencia de aparición de un término t_i en un documento d_j).

1. Caching: Una de la técnicas más utilizadas para aumentar la performance en los sistemas de búsqueda a gran escala es el *caching*, que se basa en la idea fundamental de almacenar en una memoria de rápido acceso los ítems que volverán a aparecer en un futuro cercano. En un motor de búsquedas esto se realiza en varios niveles, por ejemplo: resultados [23], *posting lists* [33], intersecciones [18, 14] y documentos [27].

Actualmente, nuestro grupo se enfoca principalmente en el problema de *caching* de resultados, *posting lists* e intersecciones. En el primero de los casos, se está trabajando con políticas de admisión para cachés de resultados utilizando aproximaciones basadas en *streaming*, ya que las consultas arriban en flujo y a altas tasas. La política de admisión se modela como un problema de clasificación y se lo aborda usando árboles de decisión dinámicos (Hoefding Adaptive Trees [3]) que permiten adaptarse a los patrones de consultas que evolucionan en el tiempo. Complementariamente, se está estudiando el almacenamiento de cantidades variables de resultados en caché, de acuerdo a la consulta.

Por otro lado, en cuanto al problema del *caching* de intersecciones, se está trabajando en el diseño

de políticas de reemplazo dinámicas para estructuras de datos que originalmente estaban definidas de forma estática como el caché Integrado de Listas + Intersecciones [28]. De esta manera, se pretende adaptar la estructura de datos y sus algoritmos de gestión para soportar reemplazo dinámico. Se evalúa tanto la tasa de aciertos (*Hit Rate*) como el costo que permite ahorrar este tipo de caché de forma que impacte positivamente en las prestaciones.

En cuanto al caché de posting lists, se intenta mejorar el uso del espacio en memoria. A partir de que los índices invertidos se almacenan en bloques comprimidos se está diseñando un esquema que permite cachear solo “algunos” bloques, de acuerdo a diferentes criterios.

2. Procesamiento de Top-k:

A través de los años, se han desarrollado múltiples estrategias para procesar *queries* de la forma más eficiente posible. En general, se las agrupan en dos categorías: i) estrategias TAAT (*term-at-a-time*), en las que se procesa de a un término del query a la vez, y ii) estrategias DAAT (*document-at-a-time*), en las que el procesamiento se realiza de a un documento por vez. Un caso práctico es retornar los k documentos más relevantes, de acuerdo a un *score* (top- k).

Realizar esto de forma exhaustiva puede resultar costoso, por lo que se han propuesto diferentes algoritmos que buscan evitar procesar aquellos documentos que probablemente no formen parte de los k primeros resultados. Dos de los más conocidos, y que forman parte del estado del arte, son *MaxScore* [29, 13] y WAND [12], los cuales almacenan de forma global para cada *posting list* su *score* más alto (*upper bound*), y utilizan esta información para determinar si un documento candidato tiene la posibilidad de ingresar al conjunto de respuestas. Por otro lado, recientemente se ha desarrollado *Waves* [7], un algoritmo que se basa en el uso de un índice multicapa, en el cual cada capa contiene las entradas con mayor impacto de cada término.

En este trabajo se propone una extensión a *MaxScore*, utilizando no sólo uno, sino una serie de *upper bounds* para decidir si un documento puede formar parte de los top- k resultados. En esta nueva versión se debe mantener una estructura adicional similar a una *skip list* con su correspondiente *upper bound*. De esta forma, se le provee al algoritmo de más información que puede utilizar en cada iteración para evitar procesar documentos innecesariamente.

3. Estructuras Escalables: Los sistemas de búsqueda que ingestan información en tiempo real (por ejemplo, sitios de microblogging como Twitter) y la ponen a disposición (indexan) en un intervalo muy corto de tiempo requieren de enfoques diferentes a los tradicionales para la gestión de sus estructuras de datos [4, 1]. Para poder mantener la eficiencia conforme se incrementa la cantidad de información que consumen, resulta indispensable mantener el índice invertido en memoria principal y gestionarlo racionalmente, manteniendo solamente aquella información que permita alcanzar prestaciones de efectividad aceptables [6]. En este sentido, el grupo trabaja en técnicas de control del crecimiento de las estructuras de datos [26] a partir de estrategias de invalidación de entradas en el vocabulario, entre otras. Siguiendo esta línea, se propone el desarrollo de una nueva familia de algoritmos de invalidación y poda selectiva [21] de las estructuras de datos a partir de la evolución y dinámica del vocabulario.

4. Estructuras Comprimidas: La compresión de datos es una técnica bien establecida en el ámbito de los sistemas de búsqueda de gran escala y ha sido extensamente estudiada. En esta línea se trata, básicamente, de la compresión de números enteros (identificadores de documentos y frecuencias de los términos). En trabajos recientes [22], se ha mostrado que combinando diferentes técnicas se puede alcanzar mejores tasas de compresión sin afectar el tiempo de procesamiento. En nuestro grupo, se investiga el concepto de representación y compresión híbrida de una *posting list*, es decir, se combinan arreglos de bits (*bitvectors*) con listas de enteros (representación tradicional) a partir de explotar las repeticiones de patrones que aparecen en las *posting lists*. Luego, tanto los *bitvectors* como las *posting lists* resultantes se comprimen con las técnicas que mejor *tradeoff* (tasa de compresión/tiempo de descompresión) ofrecen para el problema planteado [33].

b. Procesamiento y Análisis de Grandes Datos

Las plataformas de procesamiento distribuido para Grandes Datos llevan varios años de desarrollo y utilización. Éstas proporcionan interfaces que ofrecen un nivel de abstracción considerable en la utilización de un cluster a cambio de agregar capas de software que gestionan los recursos. Ejemplos de esto son plataformas como Hadoop [30], Spark [31] y Storm. El grupo investiga cómo utilizarlas eficien-

temente en los problemas antes mencionados, principalmente bajo restricción de recursos (*commodity hardware*) [8].

Un caso puntual, es la construcción de índices mediante procesos distribuidos en un cluster, que suele ser un requerimiento para las máquinas de búsqueda. Esto se debe a que los documentos pueden ser procesados en paralelo por diferentes nodos y realizarse posteriormente un proceso de integración de los datos para la construcción del índice invertido final, el cual puede estar particionado o no. Por otro lado, aparecen estructuras de datos para índices que ofrecen un mejor rendimiento para la recuperación (bajo ciertas condiciones) como es el caso de Block-Max [9] y que son de interés cuando se requiere soportar búsquedas a escala masiva.

Otro caso abordado es el procesamiento de flujos de *streams* de vídeo provenientes de cámaras de vigilancia. Se trabaja en estudiar la escalabilidad y eficiencia de un cluster Spark [32] para el problema de monitoreo en tiempo real y la detección de figuras humanas en las secuencias de imágenes (*frames*). Esto último se modela como un problema de clasificación. La carga de trabajo se distribuye entre los nodos bajo diferentes arquitecturas y se analizan los requerimientos de hardware para cumplir con las restricciones temporales. Resultados preliminares muestran que los requerimientos para procesar el video en tiempo real utilizando las técnicas mencionadas son altos y se requiere de la optimización tanto del cluster (selección de parámetros, lo que no es trivial con este tipo de herramientas [17]) y de las piezas de software. También se trabaja con esquemas de muestreo para evitar procesar todos los *frames* del video, pero manteniendo las prestaciones.

En esta línea de investigación se utilizan plataformas y técnicas del ámbito de Grandes Datos (Por ejemplo, Hadoop/MapReduce) para estudiar, diseñar y evaluar algoritmos para diferentes problemas como los mencionados y tratar de establecer los parámetros que determinan la eficiencia del proceso, tales como propiedades de los datos de entrada o las características de un cluster particular.

Resultados y objetivos

El objetivo principal del proyecto es estudiar, desarrollar, aplicar, validar y transferir modelos, algoritmos y técnicas que permitan construir herra-

mientas y/o arquitecturas para abordar algunas de las problemáticas relacionadas con las búsquedas a gran escala y el procesamiento de grandes datos.

Se estudian problemas relacionados con estructuras de datos y algoritmos, combinándolas con técnicas de aprendizaje automático y optimización para aplicaciones de búsqueda. Se proponen mejoras que apuntan a la eficiencia en una tarea. En particular, se espera alcanzar los siguientes objetivos:

- Definir y evaluar estructuras de datos híbridas que permitan ahorrar espacio mientras mantienen la performance del sistema, amortiguando el impacto del crecimiento en la cantidad de información que se debe manejar.
- Mejorar técnicas de caching específicas en motores de búsqueda, tanto políticas de reemplazo como de admisión (tema que no ha tenido suficiente desarrollo aún). En especial, se incorpora el análisis del flujo de consultas en *streaming* y se utilizan algoritmos apropiados.
- Diseñar y evaluar versiones optimizadas de algoritmos de procesamiento de *queries*, que permitan mejorar las prestaciones de los servicios de búsqueda.
- Diseñar arquitecturas para aplicaciones específicas del área de Big Data, orientadas, por un lado, a la indexación masiva distribuida y, por el otro, al procesamiento de flujos de datos en *streaming* (como los flujos de video en tiempo real).

Formación de Recursos Humanos

Este proyecto brinda un marco para que docentes auxiliares y estudiantes lleven a cabo tareas de investigación y se desarrollen en el ámbito académico. Recientemente, se ha finalizado una tesis de la maestría en “Exploración de Datos y Descubrimiento de Conocimiento”, DC, FCEyN, UBA y tres de Licenciatura en Sistemas de Información (UNLu).

Actualmente, se están dirigiendo tres trabajos finales de la Lic. en Sistemas de Información (UNLu), hay dos pasantes alumnos y un becario CIN (Beca Estímulo a las Vocaciones Científicas). Se espera dirigir al menos dos estudiantes más por año y presentar dos candidatos a becas de investigación.

Referencias

- [1] N. Asadi, J. Lin, and M. Busch. Dynamic memory allocation policies for postings in real-time twitter search. *CoRR*, abs/1302.5302, 2013.
- [2] R. A. Baeza-Yates and B. A. Ribeiro-Neto. *Modern Information Retrieval - The concepts and technology behind search, 2nd ed.* Pearson Education Ltd., 2011.
- [3] A. Bifet and R. Gavaldà. Adaptive learning from evolving data streams. In N. M. Adams, C. Robardet, A. Siebes, and J.-F. Boulicaut, editors, *Advances in Intelligent Data Analysis VIII*, pages 249–260, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.
- [4] M. Busch, K. Gade, B. Larson, P. Lok, S. Lucenbill, and J. Lin. Earlybird: Real-time search at twitter. In *Proc. of the 28th International Conference on Data Engineering, ICDE '12*. IEEE Computer Society, 2012.
- [5] B. B. Cambazoglu and R. A. Baeza-Yates. Scalability and efficiency challenges in large-scale web search engines. In *Proc. of the Eighth ACM International Conference on Web Search and Data Mining, WSDM*, 2015.
- [6] C. Chen, F. Li, B. C. Ooi, and S. Wu. Ti: An efficient indexing mechanism for real-time search on tweets. In *Proc. of the 2011 ACM SIGMOD International Conference on Management of Data*. ACM, 2011.
- [7] C. Daoud, E. Moura, D. Fernandes, A. Silva, C. Rossi, and A. Carvalho. Waves: a fast multi-tier top-k query processing algorithm. 20:1–25, 02 2017.
- [8] T. Delvechio and G. H. Tolosa. Indexación distribuida con restricción de recursos. In *Simposio Argentino de GRANdes Datos (AGRANDA)-JAIIO 46 (Córdoba, 2017)*, 2017.
- [9] S. Ding and T. Suel. Faster top-k document retrieval using block-max indexes. In *Proc. of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '11*. ACM, 2011.
- [10] P. A. Dreyer Jr. and F. S. Roberts. Irreversible k -threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion. *Discrete Applied Mathematics*, 2009.
- [11] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, New York, NY, USA, 2010.
- [12] A. B. et al. Efficient query evaluation using a two-level retrieval process. *Proc. Proceedings of the twelfth international conference on Information and knowledge management, ACM*, pages 426–434, 2003.
- [13] T. S. et al. Optimization strategies for complex queries. *Proc. Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 219–225, 2005.
- [14] E. Feuerstein and G. Tolosa. Cost-aware intersection caching and processing strategies for in-memory inverted indexes. In *In Proc. of 11th Workshop on Large-scale and Distributed Systems for Information Retrieval, LSDS-IR'14*, 2014.
- [15] J. Gantz and D. Reinsel. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the future.*, pages 1–16., 2012.
- [16] X. Guo. Application of meteorological big data. In *Communications and Information Technologies (ISCIT), 2016 16th International Symposium on*, pages 273–279. IEEE, 2016.
- [17] H. Karau and R. Warren. *High Performance Spark: Best Practices for Scaling and Optimizing Apache Spark*. O'Reilly Media, Inc., 2017.
- [18] X. Long and T. Suel. Three-level caching for efficient query processing in large web search engines. In *Proc. of the 14th international conference on World Wide Web*. ACM, 2005.
- [19] S. Madden. From databases to big data. *IEEE Internet Computing*, 16(3):4–6, 2012.
- [20] J. Magnusson. Social network analysis utilizing big data technology, 2012.

- [21] A. Ntoulas and J. Cho. Pruning policies for two-tiered inverted index with correctness guarantee. In *Proc. of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2007.
- [22] G. Ottaviano, N. Tonello, and R. Venturini. Optimal space-time tradeoffs for inverted indexes. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, WSDM '15*, pages 47–56, New York, NY, USA, 2015. ACM.
- [23] R. Ozcan, I. S. Altıngövdü, and O. Ulusoy. Cost-aware strategies for query result caching in web search engines. *ACM Trans. Web*, 5(2), May 2011.
- [24] A. O’Driscoll, J. Daugelaite, and R. D. Sleator. ‘big data’, hadoop and cloud computing in genomics. *Journal of biomedical informatics*, 46(5):774–781, 2013.
- [25] A. Rajaraman and J. D. Ullman. *Mining of Massive Datasets*. Cambridge University Press, New York, NY, USA, 2011.
- [26] E. Rissola and G. Tolosa. Improving real time search performance using inverted index entries invalidation strategies. *Journal of Computer Science & Technology*, 16(1), 2016. ISSN: 1666-6038.
- [27] T. Strohman and W. B. Croft. Efficient document retrieval in main memory. In *SIGIR 2007: Proc. of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2007.
- [28] G. Tolosa, L. Becchetti, E. Feuerstein, and A. Marchetti-Spaccamela. Performance improvements for search systems using an integrated cache of lists+intersections. In E. Moura and M. Crochemore, editors, *String Processing and Information Retrieval*, pages 227–235, Cham, 2014. Springer International Publishing.
- [29] H. R. Turtle and J. Flood. Query evaluation: Strategies and optimizations. *Inf. Process. Manage.*, 31(6):831–850, 1995.
- [30] T. White. *Hadoop: The Definitive Guide*. O’Reilly Media, Inc., 1st edition, 2009.
- [31] M. Zaharia, M. Chowdhury, M. J. Franklin, S. Shenker, and I. Stoica. Spark : Cluster Computing with Working Sets. *HotCloud’10 Proc. of the 2nd USENIX conference on Hot topics in cloud computing*, page 10, 2010.
- [32] M. Zaharia, M. Chowdhury, M. J. Franklin, S. Shenker, and I. Stoica. Spark: Cluster computing with working sets. *HotCloud*, 10(10-10):95, 2010.
- [33] J. Zhang, X. Long, and T. Suel. Performance of compressed inverted list caching in search engines. In *Proc. of the 17th international conference on World Wide Web, WWW ’08*. ACM, 2008.
- [34] P. Zikopoulos, C. Eaton, et al. *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media, 2011.

Procesamiento Eficiente de Grafos Masivos para Aplicaciones en Redes Sociales

Andrés Giordano¹, Gabriel H. Tolosa¹, Santiago Bancharo¹
Juan M. Ortiz de Zarate², Esteban Feuerstein²
{agiordano, tolosoft, sbancharo}@unlu.edu.ar; {jmoz, efeurest}@dc.uba.ar

¹Departamento de Ciencias Básicas, Universidad Nacional de Luján

²Departamento de Computación, FCEyN, Universidad de Buenos Aires

Resumen

Las redes sociales digitales se han convertido sin dudas en una de las aplicaciones más populares de Internet y atraen a millones de usuarios que, de forma implícita, generan estructuras con propiedades emergentes que surgen del comportamiento global.

Existen diversos problemas interesantes a resolver como la formación de comunidades, la recomendación de enlaces y el estudio de la polarización de opiniones. En todos los casos, resulta motivador tanto el proceso de formación como el estudio de algoritmos eficientes para el procesamiento. Estos problemas se pueden abordar estudiando el grafo subyacente, el contenido de las publicaciones o combinaciones de ambos.

En este trabajo se proponen diversas líneas de investigación sobre los temas mencionados, con aplicaciones a grafos masivos y problemas reales. Se abordan tanto problemas algorítmicos en cuanto a la eficiencia como las interacciones entre usuarios y diferentes escenarios.

Palabras clave: Red social, algoritmos eficientes, comunidades, polarización, recomendación.

Contexto

Esta presentación se enmarca en los proyectos de investigación “Algoritmos Eficientes y Minería Web para Recuperación de Información a Gran Escala”, “Estudio epidemiológico y seguimiento de pacientes con enfermedad celiaca (EC) asistidos por software y técnicas avanzadas de análisis de datos” del Depto. de Cs. Básicas (UNLu) y “Modelos y herramientas algorítmicas avanzadas para redes y datos masivos” del Depto. de Computación de la FCEyN (UBA).

Introducción

Las redes sociales digitales se han convertido sin dudas en una de las aplicaciones más populares de Internet y han modificado la forma en que los usuarios interactúan e intercambian información. Estas redes atraen a millones de usuarios [4, 8, 10] que, de forma implícita, generan estructuras con propiedades emergentes [1] que surgen del comportamiento global.

En general, este tipo de redes tienen a nivel estructural una topología libre de escala (*scale-free*), que se caracteriza por una distribución muy sesgada en el grado de los nodos. Además, son autosimilares, por lo que es posible estudiar porciones de la red y extraer propiedades generales. En este escenario, es posible diseñar algoritmos eficientes para compartir y distribuir la información generada. Además, es especialmente interesante si se tiene en cuenta que la red es un ambiente altamente dinámico y de gran escala.

Sin embargo, hay que considerar que estas redes son procesos humanos y no meramente tecnológicos, por lo que además de la estructura subyacente, también es de importancia en contenido generado. Su mejor comprensión posibilita aprovechar la inteligencia colectiva para mejorar servicios (por ejemplo, sistemas de recomendación) y aplicar a nuevos escenarios.

Existen diversos problemas altamente interesantes sobre una red social. Uno de estos es la formación de comunidades, es decir, grupos de usuarios que se agrupan por algún criterio. En este caso, resulta motivador no solo el proceso de formación y sus implicancias sino, además, el estudio de algoritmos eficientes para la conformación de comunidades [24].

Esta es una tarea desafiante a gran escala en aspectos que van desde el tamaño y el tipo de interacción hasta las similitudes por contenido. Si bien existen diversos métodos para analizar y modelar este tipos de redes, la necesidad de algoritmos que combinen información estructural con las propiedades de los nodos es un requisito para un amplio espectro de potenciales aplicaciones concretas. Algunos de estos problemas tienen aplicación potencial en proyectos de colaboración abierta, en la salud (grupos de personas con patologías similares) o en el caso de catástrofes [20, 2].

Otra cuestión de interés actual en redes sociales es el estudio de la polarización de opiniones. Se ha visto ampliamente en elecciones presidenciales pero existe aplicado a muchos otros casos de los cuales no se conoce aún si su comportamiento es similar y, en caso que no lo sea, qué modelos aparecen¹.

En este trabajo se proponen diversas líneas de investigación sobre los temas mencionados, con aplicaciones a grafos masivos y problemas reales. Se abordan tanto problemas algorítmicos en cuanto a la eficiencia como las interacciones entre usuarios y la formación de comunidades en diferentes escenarios.

Líneas de I+D

En este proyecto se continúan líneas de I+D del grupo que incorporan análisis de grafos de redes sociales que permitan mejorar la calidad de algoritmos básicos y la utilidad de métricas características de estas estructuras de datos aplicadas a éstos. Se abordan problemas clásicos sobre grafos pero aplicados a gran escala. Además, se abordan problemas aplicados a situaciones concretas. En especial, las líneas de I+D principales son:

a. Estimación en Grafos Masivos

La estructura de un grafo G se define como $G = \{E, V\}$ donde V es el conjunto de nodos o vértices y E es el conjunto de aristas que los unen. La posibilidad de computar algunas métricas básicas, como por ejemplo, el cálculo de la distancia entre dos nodos arbitrarios, se ve afectada en cuanto a la eficiencia del proceso debido al tamaño de los grafos

¹Las hinchadas de fútbol y el referéndum por la independencia de Catalunya son otros dos ejemplo que generan estructuras de opiniones polarizadas.

actuales. Por ejemplo, en redes sociales digitales el número de relaciones (aristas) supera ampliamente la cantidad de usuarios (nodos)².

El problema de la distancia entre dos nodos tiene múltiple aplicaciones prácticas, por ejemplo, para el ranking en búsquedas³. Se lo define como la longitud del camino mas corto entre ellos y se vuelve casi inviable si se requiere responder en pocos milisegundos. Esta métrica es usada en numerosos algoritmos que apuntan a resolver problemas como la recomendación de links [25], agrupamiento de usuarios [5], entre otros.

El cómputo exacto de la distancia es prohibitivo para aplicaciones prácticas dado el tamaño de estas estructuras y la estimación del valor es una alternativa. La reducción del error asociado a estimaciones de estos tipos de métricas a un costo computacional bajo conforma una de las líneas de investigación de este proyecto. En nuestro enfoque se utiliza un conjunto de nodos, llamados *landmarks* [17], que se toman como referencia para estimar luego la distancia entre dos nodos arbitrarios. El problema de la selección de “buenos” landmarks, es decir, aquellos que permitan minimizar el error de estimación, es una pregunta abierta ya que existen diversos criterios a aplicar que consideran grafos de diferente tamaño, densidad y dinámica.

b. Formación de Comunidades

Si bien no existe una definición exacta de *comunidad* y una técnica específica para identificarlas, se puede decir que una comunidad esta compuesta por usuarios que comparten alguna característica en común. El interés en tener métodos que realicen esta tarea de forma efectiva surge de varias áreas como la política [14], medicina [22], etc. A través de comunidades detectadas se puede estudiar el comportamiento e interés en ciertos temas de las personas. En las redes sociales estas comunidades se pueden detectar a través de 3 técnicas básicas:

- **Análisis de la topología de la red:** este método es el mas simple y se basa solo en el grafo subyacente a la red, es decir, los usuarios y sus relaciones [18, 3, 12]. Si bien los algoritmos que aplican este enfoque son eficaces

²Twitter: una imagen de esta red social contiene 81306 nodos con 1768149 aristas (<https://snap.stanford.edu/data/egonets-Twitter.html>)

³Un caso es la red de contactos profesionales LinkedIn.

suelen agrupar usuarios que tratan de tópicos diferentes aunque densamente conectadas (carecen de alta precisión).

- **Análisis del contenido:** este enfoque explora el contenido de las publicaciones de los usuarios y no considera la información estructural de la red como lo es la densidad de las conexiones que puede existir en un conjunto de usuarios. En Twitter, por ejemplo, esto se refiere al contenido de los tweets separando texto libre de hashtags, urls y menciones [13].
- **Híbridos:** estos métodos utilizan los dos enfoques anteriores en conjunto [19, 9, 26, 23], agregan características como la similitud de contenido aplicada como peso o importancia de la relación entre un par de usuarios. Una vez generada esta estructura se aplica algún algoritmo de detección de comunidades conocido que use el peso de las aristas.

En este tema el grupo viene trabajando con un estudio puntual relacionado con el proyecto interdisciplinario mencionado en la sección Contexto, en la cual no solamente se realizan estudios médico/biológicos/químicos sino que, además, se trabaja con el tema sobre la red social Twitter. En particular, se aplican algoritmos de formación de comunidades utilizando los tres enfoques mencionados, detectando redes de usuarios con interés en la enfermedad celíaca y poder establecer autoridades en el tema y usuarios influyentes.

c. Estudio de “Jergas” y Polarización en Contenido

En esta línea de trabajo se enfoca el problema de las comunidades a partir de dos ideas diferentes a las anteriormente presentadas: uso de determinada jerga en particular y búsqueda de opiniones polarizadas (que permite identificar dos comunidades). Si bien muchos trabajos [15, 11] usan Twitter como laboratorio y a la política como contexto, no se ha enfocado el problema en la identificación de estas mediante únicamente la jerga expuesta en sus 140 (o ahora 280) caracteres. En base a la hipótesis de que como las comunidades políticas son fuertemente homofílicas y además teniendo en cuenta como esto es potenciado por las *Filter Bubbles*[16] de las redes sociales, se genera en torno a estas una jerga

particular, a través de la cual podemos identificar la pertenencia del usuario a una comunidad. Es decir, en base a como “habla”, se pretende predecir a que comunidad política pertenece.

Resultados preliminares muestran que las estructuras de los grafos generados por los distintos conjuntos de tweets presentan patrones en común que abren las puertas a nuevas preguntas que pueden ser abordadas de manera interdisciplinaria entre la computación y las ciencias sociales.

d. Clustering

Las técnicas de clustering permiten agrupar items con características similares. Existen diversas técnicas que se aplican a casos puntuales con diverso grado de eficacia. Esta línea complementa uno de los puntos anteriores ya que se puede aplicar para la formación de comunidades donde las observaciones u objetos a agrupar son los usuarios y las dimensiones de los datos están dadas por las características de los mismos extraídas de la topología y/o el contenido publicado. Un cluster detectado será considerado comunidad pero no al revés, ya que un conjunto de usuarios que conformen una comunidad no necesariamente serán agrupados por estas técnicas. En concreto se intenta establecer las características y los algoritmos a utilizar para que la detección sea tanto eficaz como eficiente por tratarse de grafos masivos.

e. Recomendación de Enlaces

La recomendación de enlaces propone buscar y sugerir a ciertos usuarios aquellos links no establecidos que tienen alguna probabilidad de ocurrir en un futuro (por algún criterio). Aplicaciones ampliamente conocidas, como Facebook y Twitter, aplican estas técnicas de forma sistemática. Algunos trabajos proponen primero la detección de comunidades [21, 6] para luego hacer recomendación de links. Esta línea de investigación propone mejorar la calidad y eficiencia de estas recomendaciones a través del análisis de las redes aplicando algoritmos de detección de comunidades, análisis de sentimiento, métricas sobre estructuras derivadas del contenido o híbridos de estas técnicas, así como también analizar el cambio estructural y de flujo de contenidos que genera la aceptación de sugerencias en estas redes.

f. Caracterización de Objetos

Las redes sociales no solo permiten la publicación de texto sino también de imágenes y videos. Además, a estas publicaciones se le agregan metadatos como la geo-localización, fecha de creación y datos del usuario como su lugar de residencia, fecha de nacimiento, etc. Algunos trabajos proponen usar el contenido, los metadatos y la estructura de la red social para detectar objetos o eventos y caracterizarlos en base al contenido compartido por los usuarios. Hotchman et. al [7] logra encontrar ciertos patrones en las métricas de las imágenes (matiz, brillo promedio, etc) tomadas en diferentes ciudades. Este tema resulta de gran interés para el grupo y forma parte de nuestras líneas de investigación. Por ejemplo, se puede extraer información sobre ciudades y sus habitantes (aspectos culturales y sociales) a partir de analizar publicaciones de fotografías.

Resultados y Objetivos

El objetivo principal de la propuesta es estudiar, desarrollar, aplicar, validar y transferir modelos, algoritmos y técnicas que operen eficaz y eficientemente sobre grafos masivos (redes sociales digitales). Se propone profundizar sobre el estado del arte y definir, analizar y evaluar nuevos enfoques incorporando las técnicas de minería de datos. Específicamente,

- Definir nuevas estrategias de selección de nodos landmarks para la estimación de distancia entre nodos cuyo cómputo sea rápido y requiera poco espacio de información extra. Una cuestión a tener en cuenta es poder relacionar diferentes métricas con propiedades estructurales de diferentes grafos y utilizar las apropiadas en cada caso.
- Diseñar técnicas de formación de comunidades que apliquen eficaz y eficientemente a problemas concretos (por ejemplo, la comunidad de celíacos, como se mencionó), a través de enfoques híbridos principalmente.
- Definir modelos de detección de comunidades con opiniones polarizadas a partir del estudio de las jergas que se establecen alrededor de un tema o de un conjunto particular de usuarios.
- Estudiar y adaptar los algoritmos de clustering al problema anterior, con énfasis en la

performance para escalar a grafos masivos.

- Estudiar y adaptar técnicas de recomendación de enlaces y analizar el comportamiento estructural de una red bajo ciertas configuraciones de aceptación de las recomendaciones.
- Comenzar a explorar el problema de caracterización de ciudades/ciudadanos con técnicas basadas en redes sociales de imágenes. El énfasis inicial es realizar una prueba de concepto con ciudades de Argentina.

Formación de Recursos Humanos

Este proyecto brinda un marco para que algunos docentes auxiliares y estudiantes lleven a cabo tareas de investigación y se desarrollen en el ámbito académico. Actualmente, se están dirigiendo cuatro trabajos finales correspondientes a la Lic. en Sistemas de Información (UNLu) y un estudiante de doctorado del Departamento de Computación, FCEyN (UBA). Además, hay dos pasantes alumnos y un becario CIN (Beca de Estímulo a las Vocaciones Científicas). Se espera dirigir al menos dos estudiantes más por año y presentar dos candidatos a becas de investigación.

Referencias

- [1] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74, 2002.
- [2] A. Anagnostopoulos, L. Becchetti, C. Castillo, A. Gionis, and S. Leonardi. Online team formation in social networks. In *Proc. of the 21st International Conference on World Wide Web, WWW '12*, New York, NY, USA, 2012. ACM.
- [3] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008.
- [4] D. M. Boyd and N. B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 2007.
- [5] J. Edachery, A. Sen, and F. J. Brandenburg. Graph clustering using distance-k cliques. In J. Kratochvíl, editor, *Graph Drawing*, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg.
- [6] D. Feltoni Gurini, F. Gasparetti, A. Micarelli, and G. Sansonetti. Enhancing social recommendation

- with sentiment communities. In J. Wang, W. Cellary, D. Wang, H. Wang, S.-C. Chen, T. Li, and Y. Zhang, editors, *Web Information Systems Engineering – WISE 2015*, pages 308–315, Cham, 2015. Springer International Publishing.
- [7] N. Hochman and L. Manovich. Zooming into an instagram city: Reading the local through social media. *First Monday*, 18(7), 2013.
- [8] A. Java, X. Song, T. Finin, and B. Tseng. Why we twitter: Understanding microblogging usage and communities. In *Proc. of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, WebKDD/SNA-KDD '07, New York, NY, USA, 2007. ACM.
- [9] M. N. Kewalramani. *COMMUNITY DETECTION IN TWITTER.pdf*. PhD thesis, University of Maryland Baltimore County, 2011.
- [10] H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? In *Proc. of the 19th International Conference on World Wide Web*, WWW '10, 2010.
- [11] K. H. Lim and A. Datta. Following the follower: Detecting communities with common interests on twitter. In *Proceedings of the 23rd ACM Conference on Hypertext and Social Media*, HT '12, pages 317–318, New York, NY, USA, 2012. ACM.
- [12] K. H. Lim and A. Datta. A topological approach for detecting twitter communities with common interests. In M. Atzmueller, A. Chin, D. Helic, and A. Hotho, editors, *Ubiquitous Social Media Analysis*, pages 23–43, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
- [13] K. H. Lim and A. Datta. An interaction-based approach to detecting highly interactive twitter communities using tweeting links. *Web Intelligence*, 14(1):1–15, 2016.
- [14] M. Ozer, N. Kim, and H. Davulcu. Community detection in political twitter networks using non-negative matrix factorization methods. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASO-NAM)*, pages 81–88, Aug 2016.
- [15] M. Ozer, N. Kim, and H. Davulcu. Community detection in political twitter networks using non-negative matrix factorization methods. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASO-NAM)*, pages 81–88, Aug 2016.
- [16] E. Pariser. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Group , The, 2011.
- [17] M. Potamias, F. Bonchi, C. Castillo, and A. Gionis. Fast shortest path distance estimation in large networks. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, CIKM '09, pages 867–876, New York, NY, USA, 2009. ACM.
- [18] M. Rosvall, D. Axelsson, and C. T. Bergstrom. The map equation. *The European Physical Journal Special Topics*, 178(1):13–23, Nov 2009.
- [19] Y. Ruan, D. Fuhry, and S. Parthasarathy. Efficient community detection in large networks using content and links. *CoRR*, abs/1212.0146, 2012.
- [20] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: Real-time event detection by social sensors. In *Proc. of the 19th International Conference on World Wide Web*, WWW '10, New York, NY, USA, 2010. ACM.
- [21] S. Soundarajan and J. Hopcroft. Using community information to improve the precision of link prediction methods. In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12 Companion, pages 607–608, New York, NY, USA, 2012. ACM.
- [22] D. Surian, Q. D. Nguyen, G. Kennedy, M. Johnson, E. Coiera, and G. A. Dunn. Characterizing twitter discussions about hpv vaccines using topic modeling and community detection. *J Med Internet Res*, 18(8):e232, Aug 2016.
- [23] E. Vathi, G. Siolas, and A. Stafylopatis. Mining and categorizing interesting topics in twitter communities. *Journal of Intelligent and Fuzzy Systems*, 32(2):1265–1275, 2017.
- [24] M. Wang, C. Wang, J. X. Yu, and J. Zhang. Community detection in social networks: An in-depth benchmarking study with a procedure-oriented framework. *Proc. VLDB Endow.*, 8(10), 2015.
- [25] Y. Zhang and J. Pang. Distance and friendship: A distance-based model for link prediction in social networks. In R. Cheng, B. Cui, Z. Zhang, R. Cai, and J. Xu, editors, *Web Technologies and Applications*, Cham, 2015. Springer International Publishing.
- [26] Y. Zhang, Y. Wu, and Q. Yang. Community discovery in twitter based on user interests. *Journal of Computational Information Systems*, page 2012.

NoSQL: MODELOS DE DATOS Y SISTEMAS DE GESTIÓN DE BASES DE DATOS

Silvina Migani¹, Cristina Vera¹, María Inés Lund²

¹Departamento de Informática - ²Instituto de Informática, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de San Juan
silvina.migani@gmail.com; civerados@gmail.com; mlund@iinfo.unsj.edu.ar

RESUMEN

Los Sistemas de Gestión de Bases de Datos (SGBDs) NoSQL surgieron como una alternativa de solución a problemas no resueltos eficientemente por los SGBDs tradicionales. Sin embargo, a diferencia de las bases de datos relacionales, el término NoSQL comprende múltiples modelos de datos, cada uno con características diferentes. Además, hoy en día existen muchos SGBDs para cada modelo de esta nueva familia de bases de datos. Por tal motivo, se considera valioso profundizar en el estudio de este nuevo paradigma, haciendo hincapié en la comparación cualitativa de los diferentes modelos y también cuantitativa de distintos gestores de bases de datos NoSQL a través de un benchmark que establezca puntos de referencia. Así, los datos generados por este estudio permitirán orientar a los usuarios en la elección del SGBD apropiado ante un problema específico.

CONTEXTO

Este trabajo forma parte del proyecto “NoSQL: UN NUEVO PARADIGMA EN BASES DE DATOS”, de carácter bi-anual, presentado en convocatoria de la UNSJ en Diciembre de 2017. Se encuentra, al momento de esta presentación, en evaluación por pares externos y se espera su aprobación.

El proyecto se encuentra estrechamente relacionado con las dos asignaturas de Bases de Datos de las carreras de Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información de la Universidad Nacional de San Juan (Base de Datos y Tópicos Avanzados de Base de Datos) y con la asignatura Ingeniería de Software II (también de ambas carreras). El equipo de trabajo está formado por docentes investigadores del Departamento y del Instituto de Informática. Esta conjugación fortalecerá fundamentalmente el análisis y la comparación de los diferentes modelos existentes dentro de este nuevo paradigma.

1. INTRODUCCIÓN

Base de Datos NoSQL

Carlo Strozzi en 1998, utilizó por primera vez la expresión NoSQL para referirse a una base de datos open-source relacional, que prescindía del lenguaje SQL [1]. Claramente no es el significado que en la actualidad se le da a dicho término. Luego, Eric Evans, en 2009, reintrodujo el término para referirse a bases de datos no relacionales, distribuidas, linealmente escalables, de código abierto y que no garantizaban las tradicionales propiedades ACID (Atomicity – Consistency – Isolation - Durability) [2].

Actualmente, si bien se sigue discutiendo sobre la conveniencia o no del nombre, es globalmente conocido y aceptado que el término NoSQL refiere a bases de datos de esquema flexible, distribuidas y que no se ajustan al modelo tradicional de transacciones ACID [3].

En el ámbito de un simposio de “Principios de Computación Distribuida” de ACM en el año 2000, Eric Brewer enuncia el teorema de CAP, donde manifiesta que todo sistema distribuido con datos compartidos no puede satisfacer simultáneamente las tres características siguientes: Consistencia, Alta Disponibilidad y Tolerancia a Particiones [4,5,6]. Consecuentemente, los SGBDs NoSQL al no poder satisfacer en su totalidad las tres propiedades, adhieren al modelo BASE (Basically Available, Soft State, Eventually Consistent) [7]. Es decir, básicamente disponible, de estado flexible y eventualmente consistente.

Según [3] los sistemas NoSQL mantienen seis propiedades fundamentales:

1. Habilidad de escalar horizontalmente sobre muchos servidores.
2. Habilidad de replicar y distribuir datos (particiones) en muchos servidores.
3. Interfaz o protocolo simple a nivel de llamada (en contraste con el enlace de SQL).
4. Modelo de concurrencia más débil que el ACID (típico de los sistemas de bases de datos relacionales).
5. Uso eficiente de RAM e índices distribuidos.
6. Habilidad de agregar dinámicamente nuevos atributos a los registros de datos.

Varios autores [3][8] clasifican los sistemas de bases de datos NoSQL según el modelo de datos subyacente en:

- **Clave-Valor:** Sistemas que almacenan valores vinculados a un índice que se basa en una clave definida por el programador.
- **De Documentos:** Sistemas que almacenan documentos y proveen un lenguaje de consulta. Además, los documentos pueden ser indexados.
- **De Registros Extensibles u Orientados a Columnas:** Estos sistemas almacenan registros extensibles que pueden particionarse vertical u horizontalmente en diferentes nodos de una red.

Benchmarks

Desde los inicios de la informática, los benchmarks han sido ampliamente utilizados como marco de referencia para obtener resultados objetivos al comparar sistemas. En general, efectuar las pruebas no es una tarea sencilla, requiere consideraciones meticulosas e iteraciones para poder obtener resultados certeros. Actualmente, existe una amplia gama de benchmarks para evaluar el rendimiento de diferentes componentes de software y hardware y dentro del dominio de los sistemas NoSQL se han propuesto varios. La elección de un benchmark específico debiera considerar los siguientes criterios fundamentales: relevancia dentro de su dominio, simplicidad y escalabilidad [10].

La aparición de Internet generó grandes cambios en la informática. Las aplicaciones comenzaron a evolucionar y a demandar nuevos requerimientos que los SGBDs tradicionales no pudieron satisfacer adecuadamente. En consecuencia, surgieron nuevos modelos de bases de datos y gestores que los soportaron: bases de datos de documentos, clave-valor, orientados a columnas; entre los más destacados [3, 8], enunciados anteriormente. Es así que

actualmente, bajo el paraguas NoSQL, coexiste una gran variedad de sistemas de bases de datos. Por eso, se considera valioso poder identificar, distinguir y caracterizar estos diferentes enfoques; además de evaluar la performance de diferentes gestores de bases de datos. El conocimiento servirá de guía en futuros problemas a resolver en escenarios concretos.

Las actividades a desarrollar son las siguientes:

- Estudio exploratorio: Se definirá la estrategia de búsqueda y luego, ejecutará el mapeo sistemático de la literatura [8] que permita recolectar información sobre sistemas de bases de datos NoSQL y sobre los benchmarks existentes dentro de esta temática.
- Estudio y ensayo de SGBDs NoSQL: Luego de identificar algunos sistemas de gestión destacados, se procederá a estudiarlos y realizar ensayos que permitan experimentar diferencias y semejanzas.
- Selección, estudio y utilización de un benchmark apropiado para la comparación cuantitativa (tiempo de respuesta) de diferentes SGBDs NoSQL.
- Análisis de los datos recogidos, comparando cualitativa y cuantitativamente los diferentes sistemas de gestión de bases de datos NoSQL.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación seguida por el proyecto marco de este trabajo, es el estudio y comparación de modelos y gestores de bases de datos NoSQL.

3. RESULTADOS Y OBJETIVOS

Se pretende alcanzar un amplio conocimiento de conceptos relativos al paradigma NoSQL:

- Necesidades de su utilización
- Características relevantes y distintivas
- Diferenciación con las bases de datos relacionales
- Taxonomías existentes
- Caracterización de cada tipo identificado

Y, posteriormente, con el conocimiento adquirido:

- Hacer un estudio exploratorio de los SGBD NoSQL existentes.
- Identificar sistemas de gestión destacados y realizar ensayos.
- Analizar diferentes benchmarks.
- Elegir un benchmark que permita realizar una comparación cuantitativa de los sistemas de bases de datos elegidos.
- Hacer un estudio comparativo cualitativo y cuantitativo (tiempo de respuesta) de diferentes sistemas de gestión de bases de datos NoSQL.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto contribuirá a la profundización y consolidación del conocimiento de esta área temática por parte de cada uno de sus integrantes. Además, se pretende generar distintos temas para trabajos finales de grado de las carreras del Departamento y también de postgrado. De hecho, dos de las integrantes de este proyecto, durante la vigencia del mismo, realizarán su tesis de Maestría en esta temática. Además, es de interés del grupo proponer alumnos para becas de investigación.

5. BIBLIOGRAFÍA

- [1] Strozzi, C. (1998). NoSQL-A relational database management system. *Lainattu*, 5, 2014.
- [2] Evans, E. (2009). NoSQL: What's in a name. Eric Evans's Weblog. http://blog.sym-link.com/2009/10/30/nosql_whats_in_a_name.html. [Ultimo Acceso: Diciembre de 2017].
- [3] Cattell, R. (2011). Scalable SQL and NoSQL data stores. *ACM Sigmod Record*, 39(4), 12-27.
- [4] Brewer, E., & Fox, A. (1999). Harvest, yield, and scalable tolerant systems. In *HOTOS Proceedings of the The Seventh Workshop on Hot Topics in Operating Systems*.
- [5] Brewer, E. (2010). A certain freedom: Thoughts on the CAP theorem. 29th ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing PODC, 335–335.
- [6] Abadi, D. (2012). Consistency tradeoffs in modern distributed database system design: CAP is only part of the story. *Computer*, 45(2), 37-42.
- [7] Pritchett, D. (2008). Base: An acid alternative. *Queue*, 6(3), 48-55.
- [8] Makris, A., Tserpes, K., Andronikou, V., & Anagnostopoulos, D. (2016). A classification of NoSQL data stores based on key design characteristics. *Procedia Computer Science*, 97, 94-103.
- [9] Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64, 1-18.
- [10] Gray, J. (1992). *Benchmark handbook: for database and transaction processing systems*. Morgan Kaufmann Publishers Inc..

Infraestructura de Acceso a Datos Primarios con Aporte de Semántica en Repositorios Digitales

Renato Mazzanti^{1,2,3}, Marcos Zarate^{1,2,4}, Gustavo Samec^{1,2,3}, Carlos Buckle^{1,2}

¹ Departamento de Informática, Facultad de Ingeniería, Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB), Puerto Madryn, Chubut, Argentina.

² LINVI, Laboratorio de Investigación en Informática, Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB)

³ Unidad de Gestión de la Información, Centro Nacional Patagónico, Consejo Nacional de Investigaciones Científicas y Técnicas (CCT CONICET-CENPAT)

⁴ Centro para el Estudio de Sistemas Marinos, Centro Nacional Patagónico, Consejo Nacional de Investigaciones Científicas y Técnicas (CESIMAR-CENPAT-CONICET)

renato@cenpat-conicet.gob.ar, zarate@cenpat-conicet.gob.ar, gsamec@cenpat-conicet.gob.ar, cbuckle@unpata.edu.ar

Resumen

En la actualidad hay grandes retos científicos que necesitan analizar una gran cantidad y diversidad de datos, por ello, el acceso a los datos primarios producidos por las investigaciones está cobrando cada vez mayor relevancia. Compartir estos datos, brindar acceso a los mismos y permitir su reutilización genera innumerables ventajas para la comunidad científica. La Argentina cuenta con la ley 26.899 que promueve la creación de políticas y mecanismos para la gestión de datos primarios científicos a nivel nacional. Este trabajo, presenta un proyecto de investigación aplicada y desarrollo experimental que aportará modelos y componentes para las infraestructuras de Repositorios Digitales Institucionales (RDI) apuntando a facilitar el acceso a datos primarios. Hará uso de teorías y tecnologías de la web semántica para la definición de consultas integradas entre repositorios, como así también para la generación de productos de síntesis que faciliten la tarea de análisis y descubrimiento de conocimiento a los investigadores. El proyecto trabajará sobre las bases de datos de los Sistemas Nacionales de Datos Biológicos y del Mar y además pretende fortalecer el RDI de la Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB) actualmente en desarrollo.

Palabras clave: Datos Primarios, Repositorios Digitales, Web Semántica, Gestión de datos científicos.

1. Contexto

El proyecto Infraestructura de acceso a Datos Primarios con aporte de semántica en Repositorios Digitales se elaboró en el Laboratorio de Investigación en Informática (LINVI) de la UNPSJB y está integrado por docentes investigadores de la Facultad de Ingeniería Sede Puerto Madryn, con participación de estudiantes y graduados de las carreras de dicha Sede. Colaboran en el equipo de trabajo, docentes investigadores de la Universidad Nacional del Sur (UNS) y de la Universidad Nacional de San Luis (UNSL). El proyecto fue avalado por el Consejo Directivo de la Facultad de Ingeniería de la UNPSJB y será financiado por la Secretaría de Ciencia y Técnica de la Universidad para llevar a cabo durante 2018-2020. Para el desarrollo del proyecto se utilizará como repositorio documental el RDI actualmente en desarrollo en la UNPSJB. Dado que integrantes del proyecto desarrollan tareas científicas y de soporte en el Centro Científico Tecnológico CONICET-CENPAT, se tomará como referencia el RDI de dicho Centro [1] y el RDI CONICET Digital [2]. Respecto a repositorios públicos de datos primarios se utilizarán el Sistema Nacional de Datos Biológicos (SNDB) [3] y el Sistema Nacional de Datos del Mar (SNDM) [4].

2. Introducción

Recientemente se sancionó en Argentina la Ley 26.899 “Repositorios Digitales Institucionales de Acceso Abierto” que establece la obligatoriedad de desarrollar repositorios digitales, propios o compartidos, por parte de los organismos e instituciones públicas que componen el Sistema Nacional de Ciencia, Tecnología e Innovación. Esta ley requiere el establecimiento de políticas institucionales para la gestión, el acceso público y la preservación de datos primarios de investigación. En éste contexto la UNPSJB adhirió a ésta iniciativa y promovió la creación del RDI según la Ordenanza Nro. 159 del Consejo Superior.

Como ha sucedido en el resto del mundo, la mayoría de los RDIs implementados, alberga material documental, tradicionalmente asociados a la producción científica (artículos, disertaciones, reportes, libros, tesis, etc.) y un porcentaje muy bajo de datos científicos [5], lo que dificulta la posibilidad de vincular las publicaciones con los datos que le dan sustento. Las necesidades actuales de interoperabilidad, usabilidad y uniformidad de vocabularios han llevado a repensar los modelos clásicos de organización del conocimiento en RDIs. En este contexto es donde comienzan a adquirir relevancia las búsquedas semánticas, como herramienta para la interrelación inteligente entre repositorios abiertos. Incorporar contenido semántico permite que las búsquedas no sean sobre “bolsas de palabras” sino que sean búsquedas conceptuales que permitan el descubrimiento inteligente de conocimiento [6]. Por otra parte, es necesaria la producción de soluciones que acompañen y fortalezcan las normativas antes mencionadas sobre datos abiertos, en la cual gran parte de los datos científicos de investigaciones financiados con fondos públicos deben ser de dominio público y permanecer disponibles para otros proyectos que lo requieran. La posibilidad de incorporar metadatos a los conjuntos almacenados y definir modelos conceptuales permite la utilización de dichos datos a las generaciones futuras, independientemente de la

supervivencia de los integrantes del proyecto que originaron los datos [7]. Este proyecto plantea la definición de modelos y componentes para las infraestructuras de RDI tendientes a facilitar el acceso y la explotación de datos primarios residentes en diferentes bases de datos. Para ello, hará uso de teorías y tecnologías de la web semántica, abordará la vinculación entre las publicaciones y los datos (data-citation) y aspectos relacionados con la generación de productos de síntesis que faciliten la tarea de análisis y descubrimiento de conocimiento a los investigadores. Como plataforma para el desarrollo de prototipos se utilizara el RDI de la UNPSJB, actualmente en desarrollo y dado que los integrantes del proyecto se especializan en la gestión de datos primarios biológicos y del mar, se tomarán como referencia los RDIs de CONICET Digital y del CCT CONICET-CENPAT y los repositorios de datos científicos de acceso público existentes de los Sistemas Nacionales: SNDB [3] y SNDM [4], ambos son iniciativa del Ministerio de Ciencia, Tecnología e Innovación Productiva (MINCYT) conjuntamente con el Consejo Interinstitucional de Ciencia y Tecnología (CICyT) enmarcada dentro del Programa de Grandes Instrumentos y Bases de Datos.

3. Motivación

La mayor parte de las universidades y centros de investigación de todo el mundo, incluido nuestro país, disponen ya de repositorios institucionales que almacenan los resultados de investigación de sus miembros (principalmente artículos, comunicaciones a congresos y tesis doctorales), en este sentido la Ley Nacional 26.899 establece que los datos primarios de las investigaciones científicas financiadas con fondo públicos deben estar disponibles para que los usuarios de éste tipo de material puedan en forma gratuita, leer, descargar, copiar, distribuir, imprimir, buscar o enlazar los textos completos de los artículos científicos, y usarlos con propósitos legítimos ligados a la investigación científica en consonancia con el modelo de Acceso Abierto (AA) [7].

En la actualidad existen limitaciones para lograr los objetivos que promueve el AA. Un análisis crítico de la literatura disponible indica que hay limitaciones relacionadas con:

1. La calidad de los datos, ya que hay errores que podrían poner en entredicho los resultados de una investigación [8, 9].
2. Integración de los datos primarios de diferentes investigaciones dado que la forma de relacionarlos carece de vocabularios controlados o términos comunes que puedan ser relacionados.
3. Existe un alto porcentaje de la información que describe éstos datos primarios (metadatos) que suele ser semiestructurada [10]. Desde el punto de vista semántico, esto significa que puede existir un conocimiento implícito sobre éste contenido el cual no es explotado para descubrir nuevas relaciones entre ellos.

Al evaluar un conjunto de datos que es de una fuente desconocida, un usuario se basa fundamentalmente en las propiedades visibles del conjunto de datos, tales como título, editor, tamaño del conjunto de datos, etc. Los aspectos relativos a la calidad de los datos generalmente quedan fuera de las posibilidades de evaluación quedando en el terreno de la incertidumbre. Se pueden mencionar como antecedentes que en GBIF España¹ en 2007 se hicieron intentos de corregir ésta falencia mediante la implementación de un índice que permita a un usuario evaluar la calidad del conjunto de datos Darwin Core² que ellos publican y de éste modo hacer un seguimiento de las mejoras en la calidad de datos a lo largo del tiempo. Se trataba de un índice explícito simple el que fue mejorado en 2010 denominándose *Índice de Calidad Aparente* (ICA) [11]. ICA resume la calidad de los datos en las tres dimensiones para los datos de biodiversidad: taxonómica, geoespacial y temporal. Para esta temática el proyecto

¹ <http://www.gbif.es/>

² <http://rs.tdwg.org/dwc/>

tendrá en cuenta recomendaciones del Grupo de Interés de Calidad de Datos de Biodiversity Information Standards (TDWG 2016).

De acuerdo con [12] el surgimiento de la web semántica ha permitido obtener soluciones a problemas habituales en la búsqueda de información gracias a la utilización de una infraestructura común, mediante la cual, es posible compartir, procesar y transferir información de forma sencilla. Entre los principales trabajos podemos encontrar [13] en el cual los autores proponen una metodología y una implementación capaz de explotar metadatos de un repositorio implementado con DSpace³, haciendo uso de los beneficios de la web semántica para descubrir y vincular información. En el caso de [9] los autores describen los avances logrados en el desarrollo del repositorio digital Academic Commons (AC), de la Universidad de Columbia, como un repositorio interoperable, a través del uso de RDF y tecnologías de Web Semántica.

4. Líneas de Investigación, Desarrollo e Innovación

Este proyecto tiene como línea principal de investigación el Modelado Conceptual en la Web Semántica [14] para definir componentes que faciliten la integración de datos primarios y la explotación de información útil para los investigadores [15]. Además de ésta área, se abordarán otras relacionadas con la definición de infraestructuras para e-ciencia [16], plataformas para desarrollo de RDIs interoperables distribuidos e inteligencia para el descubrimiento de información.

5. Resultados esperados

El objetivo general del proyecto es definir componentes de arquitectura para ser implementados en RDIs como herramientas para facilitar el acceso unificado entre las publicaciones y los datos primarios que las

³ <http://www.dspace.org/>

sustentan y potenciar el descubrimiento de información útil para los investigadores.

Para ello, se relevará el estado actual respecto a la integración entre RDIs públicos de referencia y bases de datos científicos abiertos (particularmente datos biológicos) y se estudiarán teorías, tecnologías y recomendaciones definidas para la web semántica, para definir finalmente, modelos conceptuales de integración entre repositorios. También se abordarán estudios relacionados a las plataformas para publicación de datos primarios, a la implementación de RDIs, a los mecanismos de referencias (citas) y al descubrimiento inteligente de información científica.

Como desarrollo experimental, se utilizará el RDI de la UNPSJB, actualmente en desarrollo, sobre el cual se esperan alcanzar prototipos con características de usabilidad. Inicialmente, se realizó un relevamiento de las tecnologías disponibles para nuestra arquitectura, y se puntualizó sobre las plataformas D2RQ [17], Jena [18], OpenRefine [19] y GraphDB [20], entre otras. De este análisis, se determinó que las plataformas más convenientes para un primer prototipo son OpenRefine y GraphDB. El primero soporta extensiones para la creación de triplas RDF a partir de una gran variedad de formatos de entrada, tales como CSV, hojas de cálculo, JSON y el mismo formato RDF. Además, permite explorar y depurar los conjuntos de datos, aplicar transformaciones y definir vocabularios asociados a los diferentes campos, de una manera amigable. Por otro lado, GraphDB es un repositorio semántico que permite almacenar las tripletas generadas por OpenRefine y también trabajar con motores de inferencia y de consultas SPARQL [21] sobre estos datos estructurados.

6. Formación de recursos humanos

En éste proyecto participan seis docentes del Departamento de Informática de la Facultad de Ingeniería de la UNPSJB Sede Puerto Madryn. Dos de ellos están realizando carreras de doctorado y otras dos carreras de especialización y maestrías. Uno de los

autores de éste trabajo está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur y cuenta con beca interna doctoral del CONICET hasta el 2020. También forman parte del proyecto 1 (un) graduado de la carrera de Licenciatura en Informática y 6 (seis) alumnos del ciclo superior. Su participación tiene el objeto de introducirlos a la tarea científica y permitirles incorporar conocimientos sobre temas no desarrollados en la currícula de la carrera. En el caso de los alumnos que están próximos a graduarse, la intención del proyecto es guiarlos en el desarrollo de sus tesis en ésta rama de la disciplina. Otro aporte para la formación académica radica en la posibilidad que los alumnos puedan realizar Instancias Supervisadas de Formación en la Práctica Profesional en el marco de este proyecto.

Referencias

- [1] Repositorio Institucional CCT CONICET-CENPAT, Ministerio de Ciencia Tecnología e Innovación Productiva. Presidencia de la Nación Argentina. <http://www.repositorio.cenpat-conicet.gob.ar> Accedido: 2018-02-15.
- [2] Repositorio Institucional CONICET Digital, Ministerio de Ciencia Tecnología e Innovación Productiva. Presidencia de la Nación Argentina <http://ri.conicet.gov.ar> Accedido: 2018-02-15.
- [3] Sistema Nacional de Datos Biológicos, Ministerio de Ciencia Tecnología e Innovación Productiva. Presidencia de la Nación Argentina. <http://www.sndb.mincyt.gob.ar> Accedido: 2018-02-15
- [4] Sistema Nacional de Datos del Mar, Ministerio de Ciencia Tecnología e Innovación Productiva. Presidencia de la Nación Argentina. <http://www.datosdelmar.mincyt.gob.ar> Accedido: 2018-02-15.
- [5] Silvia Arano, Gemma Martínez, Marina Losada, Marta Villegas, Anna Casaldaliga y Nuria Bel. La comunidad, Recursos y datos primarios de la universitat pompeu fabra: los repositorios institucionales como infraestructuras científicas: estudio de caso. *Revista española de Documentación Científica*, 34(3):385–407, 2011.
- [6] D. C. Li, H. Liu, C. G. Chute, and S. R. Jonnalagadda. Towards assigning references using semantic, journal and citation relevance. In 2013 IEEE International Conference on Bioinformatics and Biomedicine, pages 499–503, Dec 2013.

- [7] Sarah CALLAGHAN. Preserving the integrity of the scientific record: data citation and linking. *Learned Publishing*, 27(5):S15–S24, 2014.
- [8] Ben R. Martin. Whither research integrity? Plagiarism, self-plagiarism and coercive citation in an age of research assessment. *Research Policy*, 42(5):1005 – 1014, 2013.
- [9] D. Fanelli. How many scientists fabricate and falsify research? a systematic review and meta-analysis of survey data. *Plos One*, 4(5):e5738, 2009-05-29 00:00:00.0.
- [10] Jung-Ran Park. Metadata quality in digital repositories: A survey of the current state of the art. *Cataloging & Classification Quarterly*, 47(3-4):213–228, 2009.
- [11] Francisco Pando. Quantifying quality: the apparent quality index”, a measure of data quality for occurrence datasets. *Biodiversity Information Science and Standards*, 1:e20533, 2017.
- [12] Rob Koper. Use of the semantic web to solve some basic problems in education. 6, 07 2004.
- [13] Dimitrios A. Koutsomitropoulos, Geor-gia D. Solomou, and Theodore S. Papatheodorou. Semantic query answering in digital repositories: Semantic search v2 for dspace. *Int. J. Metadata Semant. Ontologies*, 8(1):46–55, May 2013.
- [14] Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web. *Scientific American*, 284(5):34–43, May 2001.
- [15] Tony Hey, Stewart Tansley, and Kris-tin Tolle. The Fourth Paradigm: Data-Intensive Scientific Discovery. October 2009.
- [16] Giuseppe Andronico, Valeria Ardizzone, Roberto Barbera, Bruce Becker, Riccar-do Bruno, Antonio Calanducci, Diego Carvalho, Leandro Ciuffo, Marco Farget-ta, Emidio Giorgio, Giuseppe Rocca, Al-ber-to Masoni, Marco Paganoni, Federi-co Ruggieri, and Diego Scardaci. e-infrastructures for e-science: A global view. 9:155–184, 06 2011.
- [17] Christian Bizer and Andy Seaborne. Treating non-rdf databases as virtual rdf graphs.
- [18] Brian McBride. Jena: A semantic web toolkit. *IEEE Internet computing*, 6(6):55–59, 2002.
- [19] Ruben Verborgh and Max De Wilde. Using OpenRefine. Packt Publishing Ltd, 2013.
- [20] Ralf Hartmut Guting. Graphdb: Mode-ling and querying graphs in databases. In *VLDB*, volume 94, pages 12–15, 1994.
- [21] Eric Prud, Andy Seaborne, et al. Sparql query language for rdf. 2006.

Sistemas Recomendadores aplicados en Educación

María Emilia Charnelli^{1 3}, Laura Lanzarini², Javier Díaz¹

¹Laboratorio de Investigación en Nuevas Tecnologías Informáticas.

²Instituto de Investigación en Informática LIDI (III-LIDI).

Facultad de Informática. Universidad Nacional de La Plata.

³ Becario postgrado CONICET

mcharnelli@linti.unlp.edu.ar, laural@lidi.info.unlp.edu.ar, javierd@info.unlp.edu.ar

Resumen

La línea de investigación actual se centra en el estudio, diseño y desarrollo de nuevas técnicas de recomendación inteligente aplicadas en el ámbito educativo. El énfasis está puesto en la construcción de un Sistema Recomendador que sugiera materiales educativos de acuerdo a los intereses tanto de alumnos como docentes.

Como trabajo previo se desarrolló una variante de la técnica de filtrado colaborativo que utiliza el modelado de tópicos latentes para representar a los recursos a sugerir. Esta nueva técnica junto con lo realizado en trabajos de años anteriores serán utilizados para desarrollar un módulo de recomendación de materiales educativos.

Palabras clave: Sistemas

Recomendadores, Minería de Datos Educativa, Analítica del aprendizaje.

Contexto

Este trabajo de investigación se encuadra en el Proyecto de Incentivos acreditado "Internet del futuro: Ciudades Digitales Inclusivas, Innovadoras y Sustentables, IoT, Ciberseguridad, Espacios de Aprendizaje del Futuro" desarrollado en el Laboratorio de

Investigación en Nuevas Tecnologías Informáticas, LINTI.

Introducción

En los últimos años, las instituciones educativas se han embarcado en su propia exploración de grandes conjuntos de datos para mejorar los índices de retención y proporcionar una experiencia personalizada y de mayor calidad para los estudiantes. En este contexto han surgido dos áreas: La Minería de Datos Educativa (MDE) y la Analítica del Aprendizaje (AA) [1]. En el contexto de la AA, los Sistemas Recomendadores (SR) han demostrado ser herramientas sumamente útiles para brindar asistencia personalizada en distintas etapas de los procesos de enseñanza y aprendizaje [2]. Estos sistemas tienen la capacidad de predecir cuando un alumno particular preferirá un contenido o no de acuerdo a su perfil. Los contenidos personalizados que proporcionan los SR en el ámbito educativo incluyen: recursos educativos abiertos (REA), objetos de aprendizaje (OA), recursos web, posibles cursos a realizar, grupos de estudio, entre otros [3]. Es importante destacar que los contenidos recomendados deben

adecuarse con los objetivos educacionales definidos por los docentes. El sistema por lo tanto, debe proveer un mecanismo para que los educadores puedan controlar las sugerencias que se van a realizar. Es por esto que el algoritmo recomendador debe poseer la capacidad de adaptarse automáticamente a los objetivos educacionales de los alumnos haciéndolos coincidir con los objetivos educacionales propuestos por los docentes [4].

Sistemas Recomendadores aplicados en Educación

A partir de las líneas de investigación presentadas en WICC 2017 [5], se continuaron estudiando diferentes técnicas de recomendación inteligente, principalmente las de filtrado colaborativo basadas en el enfoque ítem-ítem. Actualmente se está desarrollando un SR que se aplicará como una extensión de un entorno virtual de enseñanza y aprendizaje, que utilizando información de las preferencias y necesidades de los alumnos y docentes podrá recomendar materiales educativos.

En [6] se propuso un nuevo método de recomendación basado en el enfoque ítem-ítem que utiliza un modelo de tópicos latentes para modelar a los artículos a partir de su información textual. Esto permitió obtener los tópicos que mejor describían a los ítems y como éstos se relacionaban entre sí. La metodología utilizada en el método propuesto y las métricas de validación aplicadas presentaron resultados preliminares satisfactorios y competitivos frente a métodos tradicionales.

Modelado de perfiles de usuarios y modelado de recursos

Otro de los puntos centrales de esta línea de investigación consiste en modelar perfiles dinámicos de los alumnos a través de información provista por los entornos virtuales de enseñanza y aprendizaje, sistemas de gestión, redes sociales, entre otros [8][9]. Un algoritmo recomendador funciona de forma precisa si dispone de perfiles de usuarios bien construidos [10].

Generalmente, los SR aplicados en el ámbito educativo se centran en los estudiantes como el principal consumidor de recursos [11], pero no consideran las preferencias de los docentes [12]. Por lo que resulta de interés que las recomendaciones se ajusten también a las necesidades de los docentes, como así también que éstos puedan controlar la calidad y los contenidos de los materiales.

Existen una gran variedad de repositorios de recursos educativos como Merlot, OER Commons, que brindan sobre cada recurso descripciones, valoraciones, comentarios y metadatos que permiten conocer: título, objetivo didáctico, competencias, tipo de material (diapositivas, apunte, objeto de aprendizaje, etc.), para qué nivel educativo está dirigido, idioma, entre otras características. Y esta información adicional puede ser utilizada por los SR. En [5] y [6] se utilizó una técnica de modelado de tópicos latentes para representar a los recursos a través de su información textual y se estableció una medida de similitud para poder compararlos a través de sus tópicos más representativos. Actualmente, se están estudiando y analizando otras técnicas de modelado de ítems, lo que permitirá al SR asistir a los docentes con los materiales que se sugieran.

Líneas de Investigación, Desarrollo e Innovación

- Estudio y análisis de algoritmos de recomendación aplicados en el ámbito educativo.
- Estudio de métricas para evaluar la performance de las recomendaciones obtenidas con diferentes algoritmos.
- Diseño e implementación de algoritmos de filtrado colaborativo y técnicas híbridas para la recomendación de materiales educativos.

Resultados y Objetivos

- Desarrollo de un algoritmo de filtrado colaborativo basado en el enfoque ítem-ítem que utiliza modelado de tópicos latentes.
- Detección y generación de características representativas de los materiales educativos a recomendar.
- Desarrollar un módulo de recomendación para sugerir materiales educativos a partir de las preferencias de los alumnos y los docentes.

Formación de Recursos Humanos

Dentro de los temas involucrados en esta línea de investigación, actualmente hay un becario de postgrado CONICET realizando su doctorado.

Referencias

[1] Romero, C., & Ventura, S. Data mining in education. *Wiley*

Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3(1), 2013, pp. 12-27.

[2] Drachsler, H., Verbert, K., Santos, O. C., & Manouselis, M. Panorama of recommender systems to support learning. En *Recommender systems handbook, 2015*, pp.421-451. Springer US.

[3] Chrysafiadi, K., & Virvou, M. Student Modeling for Personalized Education: A Review of the Literature. En *Advances in Personalized Web-Based Education*, 2015, pp. 1-24. Springer International Publishing.

[4] Huebner, R. A. (2013). A Survey of Educational Data-Mining Research. *Research in higher education journal*, 19.

[5] Charnelli, M. E.; Lanzarini, L., & Díaz, J. Personalización de la Educación a través de Sistemas Recomendadores Dinámicos. *XIX Workshop de Investigadores en Ciencias de la Computación*. 2017, Buenos Aires.

[6] Charnelli, M. E., Lanzarini, L., & Díaz, J. Recommender System Based on Latent Topics. En *Argentine Congress of Computer Science*, 2017, pp. 179-187. Springer, Cham.

[7] Charnelli, M. E., Lanzarini, L., & Díaz, J. New Item Recommendation Method Based on Latent Topic Extraction. En *Proceedings of the 3rd Conference on Business Analytics in Finance and Industry*, 2018, Chile.

- [8] Dlab, M. H., & Hoic-Bozic, N. Recommender system for web 2.0 supported elearning. *En IEEE Global Engineering Education Conference*, 2014, pp. 953-956. IEEE.
- [9] Li, Y., Zheng, Y., Kang, J., & Bao, H. Designing a Learning Recommender System by Incorporating Resource Association Analysis and Social Interaction Computing. *En State-of-the-Art and Future Directions of Smart Learning*, 2016, pp. 137-143. Springer, Singapore.
- [10] Tejada-Lorente, Á., Bernabé-Moreno, J., Porcel, C., Galindo-Moreno, P., & Herrera-Viedma, E. (2015). A Dynamic Recommender System as Reinforcement for Personalized Education by a Fuzzly Linguistic Web System. *Procedia Computer Science*, 55, 1143-1150.
- [11] Manouselis, N., Drachsler, H., Verbert, K., & Duval, E. (2012). *Recommender systems for learning*. Springer Science & Business Media
- [12] Peralta, M., Alarcon, R., Pichara, K.E., Mery, T., Cano, F., & Bozo, J.: Understanding learning resources metadata for primary and secondary education. *En IEEE Transactions on Learning Technologies*, 2017. IEEE

Benchmarking de Bases de Datos NoSQL para el almacenamiento de Modelos Semánticos

Nélida Raquel Cáceres, Ana Carolina Tolaba, Ricardo Daniel Pérez, Jairo Joel Maximiliano Quispe, Cintia Silvana Rodríguez, Iván Leandro Sandoval

Facultad de Ingeniería - Universidad Nacional de Jujuy
Ítalo Palanca 20 San Salvador de Jujuy – 0388 4221576
nrcaceres@fi.unju.edu.ar

RESUMEN

Las bases de datos tradicionales presentan limitaciones al momento de almacenar y procesar los datos debido a la cantidad (datos no sólo generados de transacciones sino desde sensores, dispositivos móviles o clics de la web) y estructura de los mismos (por ejemplo datos de redes sociales, datos espaciales o datos semánticos). La solución a éste último concepto es el surgimiento de una nueva tecnología denominada sistema de gestión de bases de datos no relacionales (NoSQL).

Los sistemas NoSQL se emplean cada vez más para el manejo de datos semánticos, es decir, modelos de datos que incluyen información semántica. Sin embargo, todavía es difícil comprender sus principales ventajas y desventajas en este contexto.

Este proyecto tiene la intención de caracterizar y comparar mediante un proceso de benchmarking bases de datos NoSQL orientadas a grafos para el almacenamiento y procesamiento de datos semánticos. El objetivo de la evaluación no es definir cuál es mejor, sino determinar aspectos comunes, características de consulta, e identificar las diferencias entre los sistemas NoSQL.

Palabras clave: Bases de Datos NoSQL, Modelo Semántico, Benchmarking.

CONTEXTO

Este trabajo es financiado por la Facultad de Ingeniería de la Universidad Nacional de Jujuy, en el contexto de la convocatoria realizada para Proyectos orientados a la Investigación Básica y a la Investigación Aplicada en el Área de Informática de la Facultad de Ingeniería. La convocatoria fue realizada con el fin de fortalecer la formación de recursos humanos en investigación. El proyecto de investigación tiene previsto una duración de un año, desde enero hasta diciembre del 2018, con la posibilidad de solicitar su extensión de acuerdo a los resultados obtenidos.

1. INTRODUCCIÓN

En la actualidad los volúmenes de datos que se generan y consumen se destacan por un crecimiento acelerado, lo que implica que los repositorios que contienen una colección de datos, no sólo sean considerados para realizar las consultas tradicionales, sino también como repositorios a partir de los cuales se puede obtener información relevante que sea útil para la toma de decisiones.

El modelo tradicional de base de datos relacional (RDBM - Relational Database Management) es consistente, y sus ventajas y desventajas son bien conocidas [1]. Sin embargo, este modelo presenta limitaciones cuando la interconectividad de los datos es importante, debido a que la manipulación de los datos en una base de datos relacional puede ser más compleja y consumir más tiempo. Esto ha llevado a emplear nuevos enfoques bajo el concepto de sistemas de bases de datos NoSQL (Not Only SQL), los cuales permiten gestionar el volumen de datos en constante aumento. Las bases de datos NoSQL son consideradas de próxima generación y se caracterizan por ser no relacionales, distribuidos, de código abierto y escalables horizontalmente [2, 3].

NoSQL corresponde a una estrategia de persistencia que no sigue el modelo de datos relacional, y que no utiliza SQL como lenguaje de consulta [4], en otras palabras, no están supeditadas a una estructura de datos en forma de tablas y relaciones entre ellas, permitiendo a los usuarios almacenar información en formatos diferentes a los tradicionales. Algunas aplicaciones de estas bases de datos pueden observarse en [5, 6, 7].

Las bases de datos NoSQL ofrecen diferentes modelos de datos como [8, 9]:

- Wide-Column Store, a diferencia de las bases de datos relacionales los nombres y el formato de las columnas puede variar de una fila a otra dentro de la misma tabla;
- Document Store, orientados a almacenar datos semiestructurados y permiten realizar consultas por atributos presentes en sus valores;
- Key-Value Stores, implementan una clave para la indexación y recuperación de datos, usualmente implementadas de manera distribuida;

- Graph, diseñadas para datos que consisten en entidades interconectadas con un número finito de relaciones entre ellos. Representan la información como un grafo usando vértices y aristas.

Las bases de datos NoSQL orientadas a grafos han cobrado mayor importancia en la actualidad debido al crecimiento en los proyectos que necesitan de una base de datos donde la importancia de la información depende de la relaciones, por ejemplo, web semántica [10], web mining [11] entre otros. Estas relaciones se implementan a través de modelos semánticos, los cuales permiten captar mejor el significado (semántica) de los datos contenidos en una base de datos. El objetivo de los modelos de datos semánticos es capturar el significado de los datos mediante la integración de conceptos relacionales con conceptos de abstracción más poderosos. El modelo semántico más empleado en los últimos años es el modelo ontológico.

Los sistemas NoSQL se emplean cada vez más para el manejo de datos semánticos. Sin embargo, todavía es difícil comprender sus principales ventajas y desventajas en este contexto.

Este proyecto tiene la intención de caracterizar y comparar mediante un proceso de benchmarking, bases de datos NoSQL orientadas a grafos para el almacenamiento y procesamiento de datos semánticos. El objetivo de la evaluación no es definir cuál es mejor, sino determinar aspectos comunes, características de consulta, e identificar las diferencias entre los sistemas NoSQL.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

En este proyecto de investigación además del estudio y profundización de los

conceptos inherentes a grafos, información semántica y su forma de almacenamiento, se propone comparar el desempeño de diferentes bases de datos NoSQL orientadas a grafos a través de un proceso de benchmarking. El objetivo de este proceso es generar conocimiento especializado en el área de modelado, almacenamiento y procesamiento de información semántica en lo referente a la representación y el uso del conocimiento en sistemas computacionales.

3. RESULTADOS Y OBJETIVOS

De acuerdo a las actividades planificadas para el desarrollo del proyecto, se espera:

- Recopilar información referente a bases de datos NoSQL orientada a grafos para determinar aspectos comunes, características de consulta y diferencias entre los sistemas NoSQL.
- Analizar la forma de empleo de las bases de datos NoSQL orientada a grafos.
- Elaborar y difundir los resultados de la experiencia a través de publicaciones y presentación en congresos y eventos.
- Fortalecer las capacidades de los recursos humanos (alumnos participantes) en actividades de investigación.
- Contribuir a la definición de futuros temas de tesis de grado para los alumnos participantes del proyecto de investigación.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto está siendo desarrollado por un equipo conformado por docentes investigadores del Grupo de Investigación

y Desarrollo en Ingeniería de Software (GIDIS) de la Facultad de Ingeniería de la Universidad Nacional de Jujuy, a continuación se detallan los responsables del proyecto:

- La Ing. Nélide Raquel Cáceres (Directora, Categoría de Investigación IV) quien coordina las actividades del proyecto y dirige a los integrantes del equipo. Actualmente realizando tesis de maestría vinculada al área de bases de datos.
- La Ing. Ana Carolina Tolaba (Codirectora, Categoría de Investigación V) y dirige a los integrantes del equipo. Actualmente realizando tesis de doctorado vinculada al área de modelado conceptual de datos a través de modelos semánticos.

Participan del proyecto alumnos avanzados de la carrera de Ingeniería Informática:

- Ricardo Daniel Pérez
- Jairo Joel Maximiliano Quispe
- Cintia Silvana Rodriguez
- Iván Leandro Sandoval

Con la realización de este proyecto de investigación se espera la consolidación de los miembros del grupo en especial de los alumnos como jóvenes investigadores. Además, el proyecto brindará un marco propicio para la iniciación de trabajos finales de grado de la carrera Ingeniería Informática.

5. BIBLIOGRAFIA

- [1] Miller, J. J. "Graph database applications and concepts with neo4j". In Proceedings of the Southern Association for Information Systems Conference, Atlanta, GA, USA. 2013. Vol. 2324, p. 36.

- [2] Sitalakshmi Venkatraman, K. F., Kaspi, S., & Venkatraman, R. "SQL Versus NoSQL Movement with Big Data Analytics". *I.J. Information Technology and Computer Science*, 2016, p. 59-66.
- [3] "NoSQL Databases," <http://nosql-database.org> Acceso: Octubre, 2017.
- [4] Arévalo, H. H. R., & Cubides, J. F. H. "Un viaje a través de bases de datos espaciales NoSQL". *Redes de ingeniería*, 2013, 4(2), pp. 57-69.
- [5] Rodríguez Pérez, A., Rodríguez Hernández, D., & Díaz Martínez, E. "Selección de Base de Datos No SQL para almacenamiento de Históricos en Sistemas de Supervisión". *Revista Cubana de Ciencias Informáticas*, 10(3), 159-170, (2016).
- [6] Martín, A., Chávez, S. B., Rodríguez, N. R., Valenzuela, A., & Murazzo, M. A. "Bases de datos NoSQL en cloud computing". In *XV Workshop de Investigadores en Ciencias de la Computación*. (2013, June).
- [7] Valenzo, M. R., Valencia, R. E. C., & Castro, J. M. M. "Integración de búsquedas de texto completo en Bases de Datos noSQL". *Revista Vínculos*, 8(1), 80-92. (2013)
- [8] Leavitt, N. "Will NoSQL Databases Live Up to Their Promise?". *IEEE Computer*, vol. 43, no. 2, pp. 12–14, 2010. Available on: <http://www.leavcom.com/pdf/NoSQL.pdf>
- [9] Sadalage, P. J., & Fowler, M. "NoSQL distilled: a brief guide to the emerging world of polyglot persistence". Pearson Education. Upper Saddle River, NJ. 2012.
- [10] Hayes and Gutierrez C., "Bipartite Graphs as Intermediate Model for RDF" in *Proceedings of the 3th International Semantic Web Conference (ISWC)*. LNCS, no. 3298. Springer-Verlag, Nov 2004, pp. 47–61.
- [11] Schenker A., Bunke H., Last M., and Kandel A., "Graph-Theoretic Techniques for Web Content Mining". *Series in Machine Perception and Artificial Intelligence*. World Scientific, 2005, vol. 62.

VISUALIZACIÓN EN CIENCIA DE DATOS

Franco Castro, Mag. Graciela Elida Beguerí, Mag. María Alejandra Malberti

Instituto de Informática / Departamento de Informática /Facultad de Ciencias Exactas
Físicas y Naturales / Universidad Nacional de San Juan

Av. Ignacio de la Roza 590 (O), Complejo Universitario "Islas Malvinas", Rivadavia, San Juan, Teléfonos:
4260353, 4260355 Fax 0264-4234980, Sitio Web: <http://www.exactas.unsj.edu.ar>
poyuyo2013, grabeda , amalberti @gmail.com

RESUMEN

El vertiginoso aumento de datos generados en los últimos años, ha servido de incentivo al desarrollo y evolución de la Ciencia de Datos. Big Data es un término aplicado a conjuntos de datos cuyo tamaño o tipo está más allá de la capacidad de las bases de datos relacionales tradicionales tanto para capturar, gestionar o procesar los datos con baja latencia. Esos datos provienen de sensores, video/audio, redes, archivos de registro, transacciones, web y redes sociales, gran parte de ellos generados en tiempo real y en gran escala. El análisis de Big Data permite a diferentes tipos de usuarios (analistas, investigadores, usuarios comerciales) tomar decisiones utilizando los datos que antes eran inaccesibles o inutilizables. Mediante el uso de técnicas avanzadas de análisis como análisis de texto, aprendizaje automático, análisis predictivo, minería de datos y estadísticas, las organizaciones pueden analizar diversas fuentes de datos no tratadas previamente para obtener nuevas ideas que les permitan tomar mejores y más rápidas decisiones. A las cuatro V, que representan las dimensiones de Big Data propuestas por IBM: Volumen, Variedad, Veracidad y Velocidad, se le suma una quinta V, o dimensión: Visualización, que hace referencia a la representación visual, comprensible de los datos. En el marco de Ciencia de Datos, esta línea de investigación propone analizar y caracterizar diferentes estrategias y herramientas de búsqueda de conocimiento para la toma de decisiones, según sus potencialidades de Visualización de Información y principios de Deep Learning. Éstas se aplicarán a conjuntos de datos

obtenidos desde diversas fuentes, en especial los disponibles bajo el nombre Open Data. De acuerdo a la naturaleza y magnitud de los datos, se considerarán variadas herramientas de software libre disponibles en el mercado, atendiendo a las potencialidades de visualización que las mismas ofrecen.

Palabras clave: Visualización, Visualización de Datos, Visualización de Información, Ciencia de Datos

1. CONTEXTO

La línea de investigación está inserta en el proyecto bianual 2018-2020 “Visualización y Deep Learning en Ciencia de Datos” que se desarrollará en el ámbito de la FCEF-UNSA. En el contexto de la misma se pretende dar continuidad a los proyectos “Minería de Datos en la Determinación de Patrones de Uso y Perfiles de Usuario”, “Búsqueda de Conocimientos en Datos Masivos” y “La Ciencia de Datos en grandes colecciones de datos” llevados adelante a partir de 2012. Considerando al usuario como el destinatario del proceso de búsqueda de conocimiento en datos, se investigará sobre aspectos de interpretación y percepción, con el propósito de clasificar las herramientas tratadas, en el marco de Visualización de Información y niveles de usuarios.

2. INTRODUCCIÓN

En la reunión patrocinada por la Division of Advanced Scientific Computing - National Science Foundation – EEUU en octubre de 1986, surgió la iniciativa de considerar a la

Visualización, como un método informático que “transforma lo simbólico en geométrico”, y que estudia además los mecanismos que permitan percibir, usar y comunicar la información visual.

Existen diversas perspectivas desde las cuales la visualización es abordada, entre las que se destacan la Perspectiva Cognitiva, que se apoya en la Psicología Cognitiva para analizar el sistema humano de la visión (percepción y procesamiento), y justificar las posibilidades de la visualización para ampliar el entendimiento; la Perspectiva Tecnológica, que incluye un amplio conjunto de técnicas o métodos provenientes de otras disciplinas, tales como la estadística, la minería de datos, y el procesamiento de imágenes, para facilitar el análisis desde aspectos cuantitativos y cualitativos y la Perspectiva Comunicacional, que considera la visualización como una ayuda eficaz para comunicar ideas.

Así como se distinguen los términos Dato, Información y Conocimiento, diversos autores proponen áreas de conocimiento o subcampos de visualización, entre los que se encuentran Visualización de Datos y Visualización de Información.

Entre las definiciones de Visualización de datos se encuentra la propuesta por Friendly y Denis (2006), que la consideran como “la ciencia de la representación visual de los datos”, prestando especial atención a los gráficos estadísticos. En general este tipo de visualización se orienta a la representación de los datos con fines exploratorios, encontrándose entre los gráficos habitualmente usados Tablas, Diagramas de Cajas y Bigotes, Gráficos de Barra, Gráficos de Línea, Gráficos Circulares, Gráficos de Dispersión, Gráficos de Burbujas, Infografías y Nubes de Palabras.

Respecto a Visualización de Información, Ltifi y otros. (2009) la definen como “Explotar las aptitudes perceptivas naturales

del usuario para comprender cualitativamente los volúmenes de información”. Por su parte Ignasi Alcalde, en su artículo Visualización de información: ¿arte o ciencia? la considera como “... la representación y presentación de datos que explota nuestra capacidad de percepción visual con el fin de ampliar el conocimiento”.

El proceso de visualización de información presentado en la figura 1 involucra la visualización de datos en sus tres primeras etapas: Datos-Información y Representación Gráfica. La cuarta etapa considera la comprensión a alcanzar, por parte del usuario, de la representación, involucrando aspectos de percepción e interpretación.

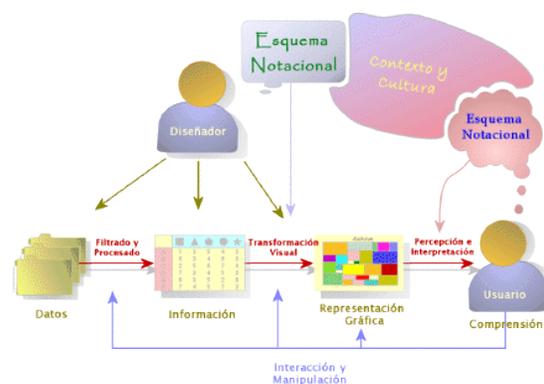


Figura 1 – Diagrama de Infovis <http://www.infovis.net/printMag.php?num=187&lang=1>

En este marco, diversos estudios sobre la percepción y cognición humana vinculados con la Visualización de Información se han realizado en los últimos años.

3. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el marco del proyecto, en lo que refiere a visualización, se pretende analizar y caracterizar diferentes estrategias de búsqueda de conocimiento en datos.

También se procura:

- Evaluar las ofertas de software libre apropiadas al área Ciencia de Datos.
- Estudiar y analizar diferentes conjuntos de datos masivos a procesar. Evaluar herramientas de software libre para arquitecturas secuenciales, paralelas y distribuidas.
- Analizar las distintas alternativas de representación de datos de entrada y determinar cuáles favorecen al desempeño de los algoritmos destinados a la búsqueda de conocimiento en datos.
- Investigar sobre aspectos de interpretación y percepción relacionados con mecanismos de visualización de información, en el contexto de búsqueda de conocimiento en datos.
- Proponer una clasificación de herramientas disponibles, en el marco de visualización de información y niveles de usuarios, y
- Caracterizar a los usuarios de acuerdo a las potencialidades de las herramientas analizadas.

4. RESULTADOS OBTENIDOS Y ESPERADOS

Atendiendo a la disponibilidad de diversas herramientas de visualización, es de interés considerar sus principales características desde el punto de vista del usuario así como capacidad de transmitir resultados surgidos de un proceso de búsqueda de conocimiento en datos.

Otro aspecto a considerar son los tipos de datos que las herramientas pueden tratar. A la hora de considerar datos, una buena opción son los datos abiertos (Open Data), esto es una filosofía y una práctica que persigue que determinados tipos de datos estén disponibles de forma libre, sin restricciones de derechos de autor, de patentes o de otros mecanismos de control (Open Knowledge Foundation). Estos conjuntos de datos se encuentran almacenados en distintos formatos. El primer desafío que enfrentamos es reconocer los formatos de uso más general y tratarlos con herramientas de software libre, en este caso

Knime, en las distintas etapas de Ciencia de Datos. El tipo de archivo depende en gran parte del tipo de dato que contiene, por ejemplo para representar datos georeferenciados, alternativas muy usadas son los KML o Shapefile. Para otro tipo de datos los formatos xls, csv, xml y json son los más utilizados. Una vez que se logra leer los distintos formatos se procede a analizar, clasificar y filtrar los datos contenidos según su tipo.

En la figura 2 se observa un ejemplo en Knime, en el cual se realiza la lectura de un archivo xml. Una vez leído el archivo es procesado con el nodo XPath, que permite separar los atributos del archivo xml en diferentes columnas.

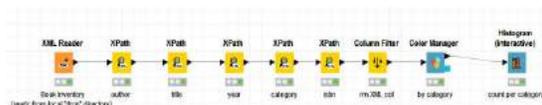


Figura 2 – Workflow de lectura y procesamiento de un archivo xml.

Finalmente la información obtenida es visualizada en un histograma como se muestra en la figura 3, respondiendo a lo que se considera Visualización de Datos.

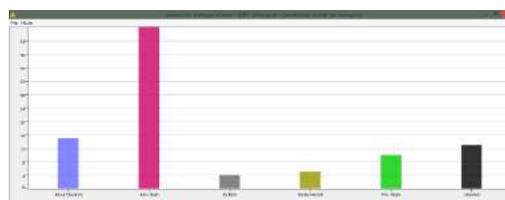


Figura 3 – Histograma representando la columna category

5. FORMACIÓN DE RECURSOS HUMANOS

La ejecución de las tareas proyectadas incidirá en una formación más profunda de los integrantes del equipo de investigación, en las recientes tecnologías de la Ciencia de Datos, en particular lo relativo a Visualización y Deep Learning. Este aspecto beneficiará de manera directa a las carreras del Departamento de Informática- UNSJ, pues las temáticas abordadas están vinculadas con las

materias en las cuales se desempeñan los integrantes de este proyecto.

Se dará continuidad a dos Becas de Investigación de Alumnos Avanzados, UNSJ, en los temas de paralelización de algoritmos y visualización, así como a dos trabajos finales de grado aplicados a diferentes conjuntos de datos y algoritmos de Deep Learning, También se prevé la generación de nuevos trabajos finales de grado para las carreras del DI, y de tesis de maestría.

6. BIBLIOGRAFÍA

- Alcalde, Ignasi. *Visualización de información: ¿arte o ciencia?* Recuperado 30/11/2017 <https://ignasialcalde.es/visualizacion-de-informacion-arte-o-ciencia/>
- Buduma, N., & Locascio, N. (2017). *Fundamentals of Deep Learning: Designing Next-generation Machine Intelligence Algorithms*. " O'Reilly Media, Inc."
- Iliinsky, N., & Steele, J. (2011). *Designing data visualizations: representing informational relationships*. " O'Reilly Media, Inc."
- Fayyad, UM, Wierse, A., y Grinstein, GG (Eds.). (2002). *Visualización de información en minería de datos y descubrimiento de conocimiento*. Morgan Kaufmann.
- Friendly, M. & Denis, D. (2006). *Milestones in the history of thematic cartography, statistical graphics, and data visualization*. Recuperado 22/11/2017- <http://web.calstatela.edu/curvebank/index/milestone.pdf>
- Karimi, H. A. (Ed.). (2014). *Big data: Techniques and technologies in geoinformatics*. CRC Press.
- Ltifi, H., Ayed, M. B., Alimi, A. M., & Lepreux, S. (2009, May). *Survey of information visualization techniques for exploitation in KDD*. In Computer Systems and Applications, 2009. AICCSA 2009. IEEE/ACS International Conference on (pp. 218-225). IEEE.
- **McCandless, D. (2010)**. *The beauty of data visualization*. John Wiley & Sons, TED website.
- Yuk, M., & Diamond, S. (2014). *Data visualization for dummies*. John Wiley & Sons.
- KNIME Analytics Platform. <https://www.knime.com/knime-analytics-platform>

ANÁLISIS DE SIMILITUD EN DOCUMENTOS DE TEXTO MEDIANTE TÉCNICAS DE CIENCIA DE DATOS BASADAS EN APRENDIZAJE PROFUNDO (DEEP LEARNING)

Calibar, Andrea Belén; Holleger, Jorge; Klenzi, Raúl Oscar
Instituto de Informática / Departamento Informática / Facultad de Ciencias Exactas Físicas
y Naturales / Universidad Nacional de San Juan
Domicilio: Av. Ignacio de la Roza 590 (O), Complejo Universitario "Islas Malvinas",
Rivadavia, San Juan, CPA: J5402DCS, 0264-260353 0264-4260355
{veneno76;wikituky;rauloscarklenzi}@gmail.com

RESUMEN

La presente propuesta pretende ingresar a un área de conocimiento actual y de creciente aplicabilidad en la extracción de conocimiento subyacente en datos de diferentes tipologías, cantidades y calidad, como es el aprendizaje profundo (Deep Learning –DL–). Aquí se propone una línea de investigación que habrá de contrastar los grados de similitud entre documentos de texto alcanzados, por medio de tres métodos y herramientas de software diferentes. Se considerará inicialmente el plugin de aprendizaje profundo *Deeplearning4J* del entorno de software libre de aprendizaje de máquina (Machine Learning –ML–) KNIME ANALYTICS 3.5.2. Una segunda alternativa a utilizar será la biblioteca GENSIM de Python, para finalmente trabajar con una versión adaptada de red recurrente creada a partir de TENSORFLOW. Se comparará el rendimiento de estas herramientas sobre datos de reservorios existentes en internet con el fin de integrarlas y explotarlas simultáneamente en entornos de hardware con CPU multinúcleos y GPU computing.

CONTEXTO

Esta propuesta se da en el marco de los proyectos CICITCA_UNSJ “Ciencia de

los Datos aplicada a grandes colecciones de datos” llevado adelante en el bienio 2016_2017 y la propuesta actual de proyecto de investigación, en evaluación, “Visualización y Deep Learning en Ciencia de los Datos”.

La elección de las herramientas a utilizar se centra esencialmente en el hecho de que KNIME es una plataforma cohesionada para científicos de datos de todos los niveles de habilidades, que proporciona un marco de ciencia de datos único y consistente. Ofrece capacidades de acceso y manipulación de datos de alta calificación, una amplia y completa gama de algoritmos y herramientas de aprendizaje automático adecuada desde principiantes hasta científicos de datos experimentados. La plataforma de KNIME se integra con otras herramientas y plataformas, como R, Python, Spark, H2O.ai, Weka, DL4J y Keras. La ayuda contextual de KNIME es más flexible que los "asistentes" fijos. La interfaz de usuario y los extensos ejemplos proporcionados con la plataforma atraen a la comunidad de científicos de datos [1].



Fig 1: Interfaz de KNIME ANALYTICS

La Fig 1 presenta la interfaz de usuario correspondiente a KNIME 3.5.2 y donde se visualizan diferentes cuadros que hacen al entorno de trabajo propiamente dicho, accesibilidad a diferentes módulos de trabajo como también ejercicios de ejemplificación, aplicación y la ayuda correspondiente a cada módulo utilizado.

1. INTRODUCCIÓN

El análisis de similitud de textos es fundamental para una amplia gama de tareas en el área de Procesamiento de Lenguaje Natural (PLN). Encontrar la similitud entre pares de textos se ha convertido en un gran reto para los especialistas del área, pudiéndose aplicar en diferentes tareas de PLN, tales como máquinas de traducción, construcción automática de resúmenes, atribución de autoría, pruebas de lectura comprensivas, recuperación de información, análisis de tendencias de investigación en el dominio académico, recomendación de cita y muchas otras, donde es prioritario medir el grado de similitud entre dos textos dados. Mayoritariamente, las métricas de similitud entre documentos han girado alrededor de similitudes sintácticas, problemática que se torna aún más compleja cuando se desea hallar la similitud semántica entre textos, los cuales en general difieren en longitud. En particular esta problemática se ve

reflejada cuando se desea encontrar el grado de similitud entre un párrafo y una sentencia, una sentencia y una frase, una frase y una palabra y una palabra y un sentido [2].

La resolución a este tipo de tareas forma parte de un gran número de problemáticas a las que el DL pretende dar solución transformándose en un área de investigación extremadamente activa que está allanando el camino para el aprendizaje de máquina moderno [3].

Dentro de la rama de la inteligencia artificial y como se aprecia en la Fig 2, el DL es un conjunto de algoritmos de *machine learning* que intenta modelar abstracciones de alto nivel en datos usando arquitecturas compuestas de transformaciones no lineales múltiples [4].

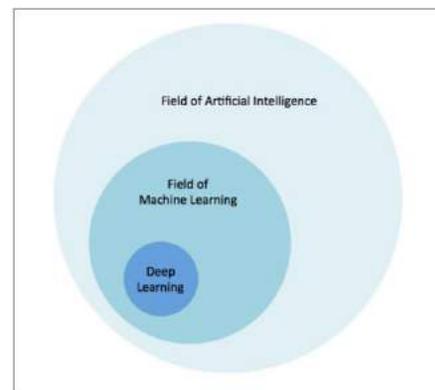


Fig 2: Deep learning como campo de Machine Learning[5].

Este tipo de aprendizaje tiene entre muchas otras características: mayor cantidad de neuronas que las redes neuronales convencionales, formas más complejas de conectar capas / neuronas en las redes, cantidad de potencia de cálculo

disponible para entrenar y extracción automática de características [5].

Una transformación previa al análisis profundo de los datos, cuando estos se presentan en forma de texto, es la incrustación de palabras (Word Embedding).

La incrustación de palabras es el nombre colectivo para un conjunto de técnicas de modelado de lenguaje y aprendizaje de características en el (PLN) donde las palabras o frases del vocabulario se asignan a vectores de números reales [9]. Cuando esto sucede, las palabras que emanan del mismo contexto o de un contexto similar pueden asociarse entre sí.

La transformación de palabras a números es necesaria porque muchos algoritmos de aprendizaje automático (incluidas las redes profundas) requieren en su entrada vectores de valores continuos, en lugar de cadenas de texto sin formato.

Una técnica de modelado de lenguaje natural de Word Embedding es word2vect, que es un grupo de modelos relacionados que se utilizan para crear incrustaciones de palabras. Estos modelos son redes neuronales que están capacitadas para reconstruir contextos lingüísticos de palabras [6].

Word2vec toma como entrada un gran corpus de texto y produce un espacio vectorial multidimensional, asignándosele a cada palabra en el corpus, un vector correspondiente en el espacio. Los vectores de palabras se ubican en el espacio vectorial de forma tal que las palabras que comparten contextos comunes en el corpus se ubican muy cerca la una de la otra en el espacio[6]

Esta y otras técnicas similares se estudiarán conformando un modelo apto para realizar el procesamiento de texto.

En la presente investigación se realizará el análisis de similitud de textos con entrenamiento de los respectivos modelos, enfocados en documentos cuyos temas están relacionados al ámbito de las ciencias de la computación y serán analizados mediante entornos aptos para utilizar técnicas de aprendizaje profundo en el procesamiento de texto.

En este contexto existen diferentes entornos de software de aprendizaje de máquina que permiten el procesamiento de dichos datos.

Para el tratamiento y análisis de los datos, se prevé utilizar, entre otras herramientas libres:

- KNIME ANALYTICS 3.5.2: Con la integración KNIME Deeplearning4J o DL4J versión 0.9.1 permite utilizar redes neuronales profundas. Se descarga de <https://www.knime.com>
- GENSIM: (librería de Python): Desarrollado originalmente por Radim Řehůřek. Gensim está bajo la licencia GNU LGPLv2.1 aprobada por OSI. Se descarga de <https://radimrehurek.com/gensim/install.html>
- TENSORFLOW: (librería de Python): desarrollado originalmente por investigadores e ingenieros del equipo Brain de Google dentro de la organización de investigación en Inteligencia Artificial de Google para realizar investigaciones de aprendizaje automático y redes neuronales profundas. Se descarga de <https://www.tensorflow.org>

En particular esta propuesta pretende una comparación de rendimiento analizando la similitud de documentos, entre el uso del módulo de aprendizaje profundo existente en la herramienta KNIME ANALYTICS 3.5.2 DL4J 0.9.1, La extensión consiste en un conjunto de nodos nuevos que permiten ensamblar de forma modular una arquitectura de red

neuronal profunda, entrenar la red con datos y utilizar la red capacitada para las predicciones. Además, es posible escribir/leer una red entrenada o sin entrenar hacia/desde un disco que permite compartir y reutilizar las redes creadas [7].

Utilizando la posibilidad de este último de extender sus capacidades mediante el uso de DL4J 0.9.1 se pretende compararla con el uso de una biblioteca de propósitos específicos como GENSIM.

GENSIM es un robusto conjunto de herramientas de modelado de temas y modelado de espacios vectoriales de código abierto implementado en Python. Utiliza NumPy, SciPy y, opcionalmente, Cython para el rendimiento. GENSIM está específicamente diseñado para manejar grandes colecciones de texto, usando transmisión de datos y algoritmos incrementales eficientes, lo que lo diferencia de la mayoría de los otros paquetes de software científico que solo se enfocan en procesamiento por lotes y en memoria [8].

Sin embargo, GENSIM no permiten utilizar el paralelismo aportado por la GPU de los actuales equipos de hardware. Aun así, ha demostrado un gran rendimiento en estudios previos. A diferencia de esta última KNIME admite GPU y esta posibilidad será explotada en la presente investigación.

La última instancia de comparación habrá de darse desde el uso de redes recurrentes contenidas en el entorno de software TENSORFLOW.

TensorFlow es una biblioteca de software de código abierto para el cálculo numérico utilizando gráficos de flujo de datos. Los nodos en el gráfico representan operaciones matemáticas, mientras que los bordes del gráfico representan los

conjuntos de datos multidimensionales (tensores) comunicados entre ellos. Su arquitectura flexible permite implementar cálculos en una o más CPU o GPU de una computadora de escritorio, servidor o dispositivo móvil con una sola API. TensorFlow fue desarrollado originalmente por investigadores e ingenieros del equipo Brain de Google dentro de la organización de investigación de Inteligencia Artificial de Google para realizar investigaciones de aprendizaje automático y redes neuronales profundas, pero el sistema es lo suficientemente general como para aplicarse en una amplia variedad de otros dominios [8].

Para efectuar el correspondiente entrenamiento, se hará uso de una gran cantidad de datos de reservorios existentes en internet y se realizará su posterior testeo con material provisto en el marco del proyecto de investigación.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el marco de la investigación se pretende:

- Analizar y describir el conjunto de estudios y prácticas requeridos en Ciencia de Datos.
- Construir y validar las diferentes instancias de modelación para la evaluación de similitudes de documentos de texto, en el mismo entorno de hardware.
- Estudiar y analizar los diferentes resultados conforme las diferentes capacidades de visualización existentes en las herramientas.

3. RESULTADOS ESPERADOS

A los efectos de la utilización del hardware y software correspondiente, se

prevé la ejecución y prueba, en forma simultánea de diferentes ejemplos de aplicación en los tres entornos de software (centrados u operados desde la interfaz de usuario de KNIME ANALYTICS) y sobre una plataforma de hardware con multiprocesamiento a nivel de CPU y GPU computing. De esta aplicación se espera encontrar mejoras en la instancia programada sobre GPU computing basada en el modelo de aprendizaje Redes Neuronales Recurrentes.

4. FORMACIÓN DE RECURSOS HUMANOS

La presente propuesta da continuidad a los proyectos que la contienen y desde allí se ha dado la posibilidad de obtener becas de investigación de alumnos avanzados actualmente en ejecución, así como la conclusión y propuestas de nuevos trabajos finales de grado y posgrado lo que sin dudas mejora la calidad y perfil de cada uno de los integrantes del grupo de trabajo. Esto permitirá el enriquecimiento de los distintos espacios curriculares de los que son responsables los diferentes integrantes del proyecto.

5. BIBLIOGRAFÍA

- [1] Gartner. **Magic Quadrant for Data Science and Machine-Learning Platforms**. Feb 2018
- [2] Darnes Villareño, Mireya Tovar, Beatriz Beltrán, Saúl León **Un modelo para detectar la similitud semántica entre textos de diferentes longitudes**. Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación, Puebla México
- [3] Nikhil Buduma. **Fundamentals of Deep Learning**-O'Reilly (2017)

[4] Y. Bengio, A. Courville, and P. Vincent., "**Representation Learning: A Review and New Perspectives**," IEEE Trans. PAMI, special issue Learning Deep Architectures, 2013.

[5] Josh Patterson and Adam Gibson **Deep Learning A Practitioner's Approach**

[6] Mikolov, Tomas; et al. "**Efficient Estimation of Word Representations in Vector Space**"

[7] <https://www.knime.com/deeplearning4j>

[8] <https://radimrehurek.com/gensim>

[9] https://es.wikipedia.org/wiki/Word_embedding

CONSTRUCCIÓN DE UNA ARQUITECTURA BI PARA OPTIMIZACIÓN DE LAS DECISIONES A PARTIR DEL ESTUDIO DEL NIVEL DE MADUREZ DE LOS ENTORNOS TECNOLÓGICOS EMPRESARIALES

Sylvia Testa¹; Lucía Rosario Malbernat²

Grupo DBI, Departamento de Sistemas, Universidad CAECE

¹Sede CABA; ²Sub sede Mar del Plata,

stesta@caece.edu.ar; lmalbernat@ucaecemp.edu.ar

RESUMEN

La transformación digital brinda la oportunidad de utilizar la tecnología para llegar a nuevos nichos de mercado, pero también plantea un enorme desafío. Las empresas se encuentran transitando el camino de sumar tecnología digital a su actual modelo, proyectando aumentar la rentabilidad proveniente de la fuerte expansión del crecimiento y alcanzar mayor eficiencia operativa.

Esta situación implica un gran riesgo para el modelo de empresa que está migrando o incorporando las nuevas tecnologías, ya que si no disponen de una arquitectura integrada, no podrán adaptarse al nuevo contexto quedando en situación desfavorable respecto al mercado, en el que la innovación tecnológica genera feroz competencia.

No obstante la tecnología por sí sola no optimiza las decisiones, es necesario instalar la “cultura digital” y concientizar a los gerentes y tomadores de decisiones sobre el potencial que reside en los datos y la necesidad de un programa que asegure la calidad, integridad y usabilidad de los datos corporativos, con equipos de tecnología y negocios trabajando en conjunto.

Se propone entonces relevar la situación actual y construir una Arquitectura de Inteligencia de Negocios para ambientes digitales (DBI: Digital Business Intelligence), que eleve el nivel de madurez tecnológica para optimizar la toma de decisiones empresariales.

Palabras clave: *Business intelligence*, arquitectura de BI, Calidad de datos, *Datawarehouse*, *Big Data*, *Analytics*, *Data*

Mining, Negocios Digitales, Innovación Empresarial

CONTEXTO

La investigación que se reporta está radicada en el Departamento de Sistemas de la Universidad CAECE, República Argentina, aprobada por R.R. 388/2017 para el período 2018-2019 con fecha de inicio 1° de abril de 2018.

Se trata de un Proyecto inter sedes ya que participarán investigadores radicados en la Sede Central y en la Subsele Mar del Plata que tendrá impacto en carreras de grado y posgrado del Departamento.

1. INTRODUCCIÓN

Los modelos de negocio son pautas esquemáticas que describen la manera en que las empresas crean y producen valor para sus clientes y el motor económico que le permite alcanzar sus objetivos, cualesquiera sean ellos, y en especial los de rentabilidad y crecimiento.

Inteligencia de Negocios o *Business Intelligence* (término acuñado alrededor de 2005 por Howard Dresner de *Gartner*, consultora internacional especializada en Tecnología de la Información y Comunicación) describe un conjunto de conceptos y métodos que mejoran la toma de decisiones utilizando información sobre hechos ya sucedidos; las tecnologías y metodologías BI permiten convertir datos en información para descubrir conocimiento.

La identidad digital se define como el conjunto de rasgos que caracterizan a un individuo o colectivo en un entorno

tecnológico. La arquitectura de BI se define como la combinación de tecnología, herramientas y procesos que permiten transformar los datos en información, esta información en conocimiento y este conocimiento dirigido a un plan o una estrategia comercial dentro del modelo de negocio.

La inteligencia de negocios es parte de la estrategia empresarial que permite optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la empresa y la capacidad de tomar buenas decisiones para así obtener mejores resultados.

El mercado de la Inteligencia de Negocios actualmente vale cientos de millones de dólares y se está expandiendo. La lista de proveedores de plataformas líderes contiene reconocidas empresas como Microsoft, IBM, SAP y Oracle, también empresas especialistas: *Tableau Software*, *QlikTech*, *Tibco Software*, *Information Builders*, etc.

Si bien parece que el éxito estaría asegurado, muchos proyectos de BI han fracasado, no tanto por la tecnología sino más bien porque la organización no supo adaptar su cultura a una nueva forma de gestionar (por falta de procesos integrados) o porque el equipo de tecnología no logró garantizar un gobierno de datos (no se trata sólo de acceder a la base de datos, es necesario comprender los metadatos y cómo se presenta la información). Estas son algunas de las causas que motivan la pérdida de confianza en los niveles decisorios, independientemente que se haya incorporado la última y más novedosa tecnología.

Estamos transitando la era Digital, la práctica del negocio está cambiando. En este nuevo escenario la única forma de tomar decisiones acertadas es cuando se dispone de información objetiva, repetible, accesible y veraz entre otros atributos, y en tiempo real. Las empresas reconocen que el camino que deben transitar es el de la innovación tecnológica que permita mejorar los indicadores de penetración e inclusión.

De hecho se está dando un gran crecimiento de empresas digitales, tal es el caso de las *Fintech* (contracción de las palabras inglesas *finance* y *technology*), Bancos digitales,

Retail digital, Portales de comercio electrónico (incluso algunos que generan negocios sin disponer de contenidos, ni stock de mercaderías, etc.).

No obstante, algunas de las preguntas que surgen son: ¿Están preparadas las empresas para asumir el costo de inversión en tecnología o el desarrollo de equipos técnicos profesionales capaces de alcanzar este objetivo? ¿Están integrados los procesos, existen bases de datos estructurados y no estructurados conectados, o son islas agregadas a los modelos tradicionales de *datawarehouse*?, ¿Existe la “Cultura de Datos” en las empresas, entienden la necesidad de datos limpios y el concepto de “calidad de datos”? ¿Entiende la alta dirección empresarial que el éxito de las decisiones tomadas depende de estas Arquitecturas BI?

Esta investigación relevará la situación actual, necesidades y aportará soluciones para acompañar a las empresas a transitar ese camino que puede implicar, además de software específico, la incorporación de almacenes corporativos del tipo *Datawarehouse* y la aplicación de técnicas tradicionales de *Data Mining* o de *Big Data*.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se ha definido como Problema a investigar el nivel de madurez de los entornos tecnológicos digitales en relación a los grandes volúmenes de datos recopilados con el fin de poder proponer una Arquitectura de BI que construya la identidad digital del cliente/usuario, mediante la integración de almacenes de bases de datos estructurados, las plataformas de almacenamiento de datos no estructurados, las herramientas de Big Data para su explotación, y las técnicas estadísticas y de Inteligencia Artificial que transforman esos datos en conocimiento, bajo un programa que implemente reglas y políticas que aseguran el gobierno de esos datos corporativos.

Se propone evaluar cómo incide ello en la toma de decisiones empresariales y en la potencialidad para nuevos negocios innovadores.

La propuesta es utilizar técnicas de I + D, de tipo cualitativas y cuantitativas, que permitan obtener resultados fiables y útiles para determinar el grado de aplicación de ésta arquitectura y su impacto en la toma de decisiones empresariales en ambiente digital, basando la investigación en la participación de los propios colectivos a investigar.

3. RESULTADOS ESPERADOS

Se ha planteado como uno de los objetivos específicos del proyecto relevar las arquitecturas BI implementadas en las empresas, especialmente en cuanto a la integración de datos estructurados y no estructurados, a la calidad de sus datos, a las herramientas con que cuentan de análisis y vinculadas a la seguridad así como también al nivel de adaptación al entorno digital.

Se espera conocer el nivel de madurez en el gobierno de datos de las empresas de modo de poder identificar y analizar los motivos y las debilidades de las empresas con menor nivel de madurez tecnológica y de calidad de datos, para evaluar el impacto en la toma de decisiones.

En base a los resultados de las actividades mencionadas anteriormente se buscará aportar insumos para la conformación de un equipo de *management*, integrado por personas de tecnología y de las áreas de negocio, cuya misión sea asegurar calidad, integridad y usabilidad de los datos corporativos para tomar decisiones más certeras; diseñar una arquitectura de BI que incluya el ecosistema de datos corporativos, plataformas (para explotación y almacenamiento de datos estructurados y no estructurados), herramientas de análisis y visualización, aplicación de técnicas estadísticas y de inteligencia artificial, interpretación de resultados y puesta en producción para ser explotadas por usuarios, alta gerencia y ejecutivos, con un objetivo final en común

que sea maximizar la eficiencia en la toma de decisiones.

Como derrame se podrá proponer un plan de acción que incluya actividades tales como la vinculación de profesionales y organizaciones, la incorporación en prácticas profesionales supervisadas a los alumnos de la Universidad, el relevamiento del mercado empresarial digital, el relevamiento de arquitectura BI en el mercado digital, la evaluación del status BI digital, el análisis de debilidades y propuesta para mejorarlo, la comunicación a las organizaciones de los resultados obtenidos, la implementación de arquitecturas BI propuestas y la evaluación de resultados mediante métricas y validaciones empíricas, lo cual permitirá recomendar correcciones y mejoras.

Se espera así, que los logros alcanzados por este Proyecto, además de generar conocimiento empresarial útil para las organizaciones que se estudien, fortalezcan la relación Universidad-Empresas, contribuyan con insumos de investigación y de práctica profesional a la Maestría en Ciencias de Datos e Innovación Empresarial e impacten positivamente en las carreras de grado Ingeniería en Sistemas y Licenciatura en Sistemas, en especial en la asignatura Inteligencia de Negocios.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación DBI (*Digital Business Intelligence*) está integrado por profesores del Departamento de Sistemas, docentes de la asignatura Inteligencia de Negocios de las carreras de grado Ingeniería y Licenciatura en Sistemas y de la Maestría en Ciencias de Datos e Innovación Empresarial, que toman como insumo académico lo producido en el marco de los Proyectos de Investigación que desarrollan.

Las investigaciones en curso del Departamento de Sistemas son propuestas cada año a los estudiantes como espacios de desarrollo de sus trabajos finales. Con este Proyecto en particular, se espera, además,

aportar organizaciones dónde los estudiantes avanzados puedan realizar sus prácticas profesionales.

5. BIBLIOGRAFIA

[CAN07] Cano, J.L. Business Intelligence: Competir con Información. Banesto, Fundación Cultur [i.e. Cultural], Madrid, España, 2007

[CS17] Chalfey D y Smith P., Digital Marketing Excellence. Publisher Routledge, version impresa y kindle, 2017.

[FOR13] Foreman J. W., Data Smart: Using Data Science to Transform Information into Insight. Editorial Wiley, 2013

[FPM92] W.J. Frawley, G. Piatetsky-Shapiro, C.J. Matheus, Knowledge Discovery in Databases: An Overview. AAAI Fall 1992, California, 1992

[JAN11] Jank W. Business Analytics for Managers. Editorial Springer, 2011.

[MA16] Marr B. y Arilla S, Big Data: la utilización del Big Data, el análisis y los parámetros Smart para tomar mejores decisiones y aumentar el rendimiento, Ediciones Teell, 2016.

[PC15] Perez M y César J., Protección de Datos y Seguridad de la Información. 4ª Edición, Editorial RA-MA, 2015.

[PEÑ02] Peña D. Análisis de Datos Multivariantes Mc Graw-Hill / Interamericana de España ISBN 9788448136109, 2002

[SC13] Sebastian-Coleman, L. Measuring Data Quality for Ongoing Improvement, Elsevier, USA, 2013

[SOA15] Soares S., Data Governance Tools: Evaluation Criteria, Big Data Governance, and Alignment with Enterprise Data Management. Editor: Mc Pr Llc, versión impresa y kindle, 2015.

[SR15] Solana A., Roca G., Big data para directivos, Editorial Empresa Activa, versión impresa y versión Kindle, 2015.

Recuperación de la Información

Ryckeboer Hugo, Sposito Osvaldo, Bossero, Julio César, Barone Miriam
Departamento de Ingeniería e Investigación Tecnológica

Universidad Nacional de La Matanza

hryckeboer@unlam.edu.ar spositto@unlam.edu.ar mbarone@unlam.edu.ar
jbossero@unlam.edu.ar

Resumen

Las Técnicas de Recuperación de Información que responden a inquietudes puntuales, se han hecho populares gracias a los buscadores ofrecidos gratuitamente a quienes recurren al Internet. Y no se hace referencia únicamente a grandes repositorios, sino también a medianos y pequeños, con miles de documentos, donde también es un inconveniente la localización del documento o los documentos que respondan a la inquietud del Usuario.

El grupo posee sus propios motores orientados a corpus estáticos. De las diversas concepciones existentes ha prestado especial atención a la indexación semántica latente conocidas por sus siglas LSI.

Una vez construidos los motores las líneas de investigación se han orientado a enfoques que permitan acelerar los mismos tanto en la búsqueda como en los preprocesos.

Uno de ellos es el uso del procesamiento paralelo, tanto en clúster de máquinas como en el uso de placas de video.

La técnica LSI es particularmente dependiente en su preproceso de un eficiente cálculo de autovalores y autovectores de matrices de gran tamaño, lo que hace incluir en nuestra temática el cálculo numérico.

Se investiga si es factible acelerar la selección de los documentos que responden a un requerimiento por medio de un particionado del corpus basándose en criterios de similitud propia de minería de datos y técnicas de selección de la parte usando redes neuronales.

En este sentido se exponen distintas líneas de trabajo a seguir, teniendo como objetivo

diseñar, implementar y probar modificaciones en los procesos de filtrado y ordenamiento de documentos, en un Sistema de Recuperación de Información (SRI), aplicando algoritmos de clustering tradicionales.

Palabras claves: Examinar, Indexar, Buscar, SRI, LSI, Minería de Datos, Agrupamiento, Algoritmos de Clasificación (Browse, Indexing, Search)

Contexto

Esta propuesta de trabajo se lleva cabo dentro de dos proyectos de investigación “*Optimización de la Recuperación de Documentos, usando como técnica base el LSI (Lematización Semántica Latente)*”, y el proyecto “*Uso de Minería de Datos para acelerar la recuperación de documento*”

Los cuales son desarrollados por el grupo de investigación del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza, en el marco de investigación del programa PROINCE.

1.Introducción

Los contenidos de los documentos digitales hoy en día, son una materia prima muy valiosa, tanto para empresas u organizaciones como para simples usuarios. Es por esto que en la Sociedad de la Información se destinan gran cantidad de recursos en almacenar grandes volúmenes de documentos, organizarlos para luego recuperar los adecuados, debido entre otras cosas, a la explosión en el número de fuentes de información disponibles en Internet que sobrepasa a toda información manual. También hay conjuntos de documentos

cerrados, por ejemplo, los legislativos, las obras de filósofos famoso, sobre los cuales se desean efectuar consultas puntuales.

Este es uno de los motivos por lo que, desde hace años, se dispone de los denominados Sistemas de Recuperación de Información (*SRI o IRS en inglés Information Retrieval Systems*), que permiten almacenar, buscar y mantener documentos, extendiendo esto a textos, imágenes, vídeos, audios y otros objetos multimedia, los cuales, utilizan técnicas de búsqueda relativas a su contenido, que son específicas para cada tipo de información.

Para evitar una dispersión en un grupo humano pequeño, los proyectos se han centrado sobre información textual en español

Las manifiestas similitudes existentes entre la recuperación de información y otras áreas vinculadas al procesamiento de la información, se repiten en el campo de los sistemas encargados de llevar a cabo esta tarea. Para Salton en [Sal86] “...*la recuperación de información se entiende mejor cuando uno recuerda que la información procesada son documentos...*”, con el fin de diferenciar a los sistemas encargados de su gestión de otro tipo de sistemas, como los gestores de bases de datos relacionales. Salton piensa que “...cualquier SRI puede ser descrito como un conjunto de ítems de información (DOCS), un conjunto de peticiones (REQS) y algún mecanismo (SIMILAR) que determine qué ítem satisfacen las necesidades de información expresadas por el usuario en la petición”

Para poder hacer aportes originales en esta temática fue fundamental tener motores completos en estado operativo.

El grupo ha privilegiado la metodología LSI porque además de su conocida habilidad para resolver ambigüedad, equivocidad y sinonimia, provee vectores descriptivos de los documentos y de consultas de menor dimensión lo que beneficia a la minería de datos.

Los motores construidos son lo suficientemente abiertos y flexibles para ser utilizado en docencia y en puestos de investigación.

Cada día se utilizan técnicas más avanzadas de análisis del contenido de los documentos con vistas a mejorar los tiempos de acceso a los documentos y la efectividad del resultado.

Por lo tanto, el problema al que se enfrenta la RI se puede definir como: “Dado un conjunto de documentos, ordenar los documentos de mayor a menor según la relevancia para una determinada necesidad ya expresada como consulta, las limitaciones perceptivas del usuario aconsejan entregar los elementos que encabezan la lista”.

Aunque una buena parte del pre proceso de organización y proceso de consultas recurren a técnicas básicas de la computación se pueden señalar algunas áreas que dan lugar a mejoras y optimizaciones las que influirán en la calidad de las prestaciones:

Análisis Textual: es una práctica ya establecida que en lugar de recurrir a la presencia o no de palabras identificadas en la consulta dentro de los documentos, es conveniente reducirlas a lexemas ignorando deflexiones propias del desarrollo del texto, pero salvando un concepto común que la palabra encierra en sus distintas formas morfológicas.

Esta actividad en sus detalles es dependiente del idioma y del campo de aplicación.

Proceso del Corpus: Los coeficientes de los vectores que describen la temática de los documentos requieren un ajuste a la luz de la totalidad de los documentos disponibles. Esto en el caso del LSI requiere hallar autovalores y autovectores de elevada dimensión, problema no trivial por los errores de redondeo del cálculo con reales. Continuamente aparecen propuestas que intentan acelerar o mejorar tales cálculos.

Resolución de las Consultas: Una forma de saber el valor de un documento frente a una consulta es enfrentar los vectores

representativos de ambos. El tiempo que insume la construcción de la lista ordenada crece con el tamaño del corpus.

Dos líneas de trabajo surgen frente a esta situación:

- Distribuir los documentos en varios procesadores
- Dividir el corpus en conjuntos más pequeños conteniendo documentos “similares” y comenzar la recuperación por la parte más promisoría.

Tanto en el dividir como elegir esta parte lo estamos encarando con técnicas de minería de datos.

Los cálculos de autovalores involucran a todos los coeficientes de la matriz lo que crea un dilema y necesidad de hallar un compromiso entre distribuir el cálculo aumentando las comunicaciones entre procesadores, o concentrarlos para no tener tales recargos de tiempos.

2. Líneas de Investigación y desarrollo

Mejorar el trabajo con RI implica, por lo que se ha dicho, considerar diferentes líneas de investigación:

- a) Acelerar la velocidad de cómputo, recurriendo a un procesamiento en paralelo
- b) Subdividir el corpus en forma inteligente de modo tal que sin gran pérdida de exhaustividad se pueda resolver la consulta examinando una o más partes de la subdivisión, excluyendo a muchas de ellas.
- c) Mejorar la lematización disponible del idioma español, dado que la misma no da resultados satisfactorios.

Respectivamente, se señalan las observaciones que serán las inquietudes centrales en estas líneas de investigación:

Los sistemas que operan en gran escala deben recurrir necesariamente al uso en paralelo de varios procesadores. Se estudia la forma de paralelizar algunos algoritmos para acelerar adecuadamente los cómputos.

Dada la posibilidad de extender la selección de documentos a corpus muy voluminosos, existen diversas ideas de subdividir el corpus en grupos aplicando técnicas de agrupamiento. Para que la subdivisión sea efectiva a los fines propuestos se debe agrupar documentos de iguales características y separar los que manifiestamente difieren. En este proyecto se pretende dominar e incorporar estas tecnologías a nuestro prototipo, con la intención de evaluar si la mejora en velocidad compensa una eventual pérdida de exhaustividad a las mismas. Las tecnologías que resuelve esto se conoce como “clustering” [Her04], de modo tal de disminuir el espacio de búsqueda cuando se procesa una consulta. El proyecto pretende subdividir un corpus en varios grupos o clúster, para realizar la búsqueda de documentos pertinentes a una consulta dada. Por lo que la hipótesis principal es que la utilización de técnicas de clustering y de aprendizaje supervisado, en un SRI, acelerará la obtención de documentos pertinentes. Los beneficios alcanzarían a los usuarios directos y a la reducción de recursos computacionales. La subdivisión encierra un riesgo de respuestas no exhaustivas que afecten a la calidad del servicio. Esto se encuentra en etapa de evaluación.

Explorar a fondo las distintas alternativas que se presentan en las diversas etapas de esta solución constituyen el centro de uno de los proyectos en curso.

Las mismas se pueden resumir en:

- ✓ Elegir técnica de clustering o incluso diseñar una nueva.
- ✓ Diseñar el algoritmo que oriente una búsqueda en particular hacia uno o varios de los subconjuntos. Las redes neuronales son una de las técnicas que se vislumbran como promisorias para esta tarea.
- ✓ Proponer técnicas de evaluación en cuanto a cobertura lograda versus tiempo de respuesta.

Pensando en el preproceso está la lematización, esta se puede desdoblar o sea repartiendo a distinto procesador, distinto documento o inclusive repartiendo párrafos. La suma de presencia de lexemas en distintos documentos también es paralelizable. La transposición de la tabla documento-termino también es realizable en paralelo diseñando cuidadosamente el algoritmo.

Objetivos Secundarios

De lo enunciado ut-supra se desprenden los siguientes objetivos secundarios:

- a) Obtención de un Corpus en español, más voluminoso que el utilizado hasta ahora, para la aplicación de los algoritmos.
- b) Evaluar y proponer algoritmos de agrupamiento para el conjunto de documentos disponibles.
- c) Ajustar los modelos predictivos de tipo supervisados, para resolver el problema de clasificación.
- d) Evaluar el alcance de los resultados a través de las métricas propuestas.
- e) Aplicar los modelos desarrollados en nuevos Corpus

3. Resultados Obtenidos/Esperados

Las técnicas de minerías de datos en uso en esta línea de investigación fueron profundizadas en proyectos previos lo que dio lugar a las siguientes publicaciones:

- 1- “Predicción del riesgo de abandono universitario utilizando métodos supervisados” En colaboración con Edwards, Diego y Pérez, Silvia (UNLaM). Trabajo presentado en el Workshop de la V Jornadas Nacionales y I Latinoamericanas de Ingreso y Permanencia en Carreras Científico – Tecnológicas. Facultad Regional Bahía Blanca. Universidad Tecnológica Nacional. Bahía Blanca. Mayo de 2016. IPECyT 2016
- 2- “Modelos de minería de datos para el diagnóstico de enfermedad de Parkinson mediante el análisis de voz”. En colaboración

con el Ing. Osvaldo Sposito, Ing. Gabriel Blanco, Mg. Mónica Giuliano y el Ing. Luis Fernández (UNLaM). Trabajo presentado en el Workshop del V Congreso Nacional de Ingeniería en Informática/Sistemas de Información. Publicación en línea - ISSN. CONAIIISI 2017. Santa Fe. Argentina.

3- “Comparación de Algoritmos de Aprendizaje Supervisado para la obtención de perfiles de alumnos desertores”. En colaboración con el Ing. Osvaldo Sposito (UNLaM). Trabajo presentado en el Workshop del IV Congreso Nacional de Ingeniería en Informática/Sistemas de Información. Publicación en línea - ISSN 2347-0372. CONAIIISI 2016. Salta. Argentina

4- “Una paralización del método de Householder” En colaboración con el Ing. Osvaldo Sposito, Ing. Hugo Ryckeboer. (UNLaM). Trabajo presentado en el XXII Congreso Argentino de Ciencias de la Computación- CACIC 2016- Universidad Nacional de San Luis San Luis.

En el marco de la línea de investigación para acelerar la velocidad de cómputo, recurriendo a un procesamiento en paralelo a lo que respecta a la lematización del idioma español disponible que no daba resultados satisfactorios, se puede concluir que con el uso de los hilos se realiza un procesamiento más rápido que con la forma secuencial.

Finalmente, respecto a los sistemas que operan en gran escala, estos deben recurrir necesariamente al uso en paralelo de varios procesadores. La utilización de los GPU, es el acelerador dominante para realizar procesamiento de cálculos en paralelo, principalmente en matrices, ello tuvo un resultado positivo: los tiempos bajan drásticamente comparando un proceso secuencial en una Pc típica de escritorio contra el procesamiento sobre cualquiera de las GPU.

En cuanto a la comparación entre las diferentes GPU, se observa que los mejores tiempos se obtuvieron para la GPU R9 390X. A medida que haya más documentos, la

diferencia de tiempos entre cada una de ellas se va notando considerablemente.

Los documentos de un Corpus son transformados en vectores descriptivos. Una consulta de usuario es también convertida en otro vector descriptivo. Para obtener un documento que satisfaga la necesidad de información del usuario, el vector de la consulta se debe enfrentar con todo el corpus, en búsqueda de similitudes. Este proceso genera un índice de relevancia, que será la salida que recibe el usuario, en forma de lista ponderada. Dividir el Corpus, de modo tal de poder desechar uno o más grupos, debería acelerar la búsqueda.

4. Formación de Recursos Humanos

Resultados en cuanto a la producción de conocimiento:

Disponer de un buen lematizador del español es una contribución al estado del conocimiento en Recuperación de Información en lengua española.

Resultados en cuanto a la difusión de resultados:

Del mismo modo que en el proyecto precedente se puso un Motor de Búsqueda a disposición de toda la comunidad académica, se hará lo mismo con el lematizador español.

Los resultados en materia de lematización y del uso de “clusters” de computadoras, serán expuestos en Congresos/Revistas.

Los lematizadores no exponen normalmente la metodología con la cual los diseñaron, con lo cual, la exposición de estos detalles, puede ser valiosa para profesionales de otras lenguas.

El armado de “clusters” obliga a resolver problemas prácticos de conexión y administración, lo que puede acortar el camino a otros investigadores que se inicien en el tema.

Los profesionales de Informática disponen de baja formación lingüística, de modo tal que su participación en este proyecto les abrió

nuevos campos de actividades: la lingüística computacional, el manejo de semántica, traducción automática. Todas estas actividades requieren un adecuado manejo morfológico del lenguaje.

Respecto a la minería de Datos se utilizaron parte de los conocimientos en 2 trabajos de tesis desarrolladas en sendas maestrías en informática:

Tesis aprobada: El Soporte Informático y su Aporte para la Inclusión Universitaria

Tesis en curso: Estudio Comparativo de Técnicas de Minería de Datos para la predicción de deserción universitaria

Resultados en cuanto a transferencia hacia las actividades de docencia y extensión:

Los sistemas de IR son eficaces en la medida que diseñen buenas estructuras de datos. Estas son estudiadas en materias intermedias de la Ingeniería de Sistemas, poder ilustrar su uso práctico, beneficiará a los estudiantes.

Del mismo modo, el uso de clusters ilustra los tópicos más avanzados de la Arquitectura de Computadoras, que también integran el Plan de Estudios.

5. Bibliografía

[Her04] “Introducción a la minería de datos”. José Hernández Orallo y otros. Editorial: Pearson. Edición: I. Año 2004

[Ven03] “Análisis Semántico Latente: una panorámica de su desarrollo”. René Venegas

Rev. signos [online]. 2003, vol.36, n.53, pp.121-138. ISSN 0718-0934. Pontificia Universidad Católica de Valparaíso. Chile

Disponible en:
<http://dx.doi.org/10.4067/S0718-09342003005300008>.

[Sal86] “Introduction to Modern Information Retrieval”. Gerard Salton, Michael J. Michael J. McGill. Ed. McGraw-Hill, Inc. New York, NY, USA. ISBN: 0070544840. 1986

[Sal68]“Automatic Information Organization and Retrieval”. Salton, G. McGraw–Hill, N.Y. 1968.

[1] PCI-Express
<http://pcisig.com/specifications/pciexpress/resources>

<https://nvlabs.github.io/moderngpu/performance.html>

[2] “*Clustering de documentos con restricciones de tamaño*”. Diego Fernando Vallejo Huanga. Trabajo Fin de Máster. Universitario en Gestión de la Información. Escola T. S. d’Enginyeria Informàtica Universitat Politècnica de València. 2015. Disponible:

<https://riunet.upv.es/bitstream/handle/10251/69089/VALLEJO->

[ClusteringdeDocumentosconRestricciones deTamaño.pdf?sequence=23](https://riunet.upv.es/bitstream/handle/10251/69089/VALLEJO-ClusteringdeDocumentosconRestriccionesdeTamaño.pdf?sequence=23)

[3] “Clusterdoc, un sistema de recuperación y recomendación de documentos basado en algoritmos de agrupamiento”. Marylin Giugni O.;Luis León G.;Joaquín Fernández.

Telematique, volumen 9 - número 2 - año 2010.

Disponible

en:<http://publicaciones.urbe.edu/index.php/telematique/article/view/913/pdf>

[4] Lindholm, Erik and Nickolls, John and Oberman, Stuart and Montrym, John,NVIDIA Tesla: A unified graphics *and computing architecture*,IEEE micro, 2008.

Estudio y análisis de técnicas de modelado de grandes volúmenes de datos jurídicos

Lilia Palomo, Norma Lesca y Juan Mulki

Facultad de Ciencias para la Innovación y el desarrollo

Universidad Católica de Santiago del Estero

lilia.palomo@ucse.edu.ar – norma.lesca@gmail.com – jmulki@ucse.edu.ar

RESUMEN

En los últimos años, a raíz de la aplicación de las tecnologías de la información, se ha producido la proliferación masiva de datos, lo que dificulta a los científicos, empresas y organizaciones mantenerse al día con los resultados y tecnologías desarrolladas, llevando a que se malgaste tiempo, esfuerzo y recursos

Actualmente se puede contar con grandes repositorios de datos disponibles a costos bajos e incluso gratuitos, listos para ser explotados por procesos informacionales.

A partir de este aumento de información, se da la necesidad de extraer de ella patrones, tendencias y/o conocimiento de forma rápida y eficiente, para lo cual los métodos tradicionales han tenido que evolucionar en busca de rendimiento y escalabilidad. El gran contenido de valor que genera este tipo de información está permitiendo a las organizaciones una mejora en la toma de sus decisiones, lo que conlleva a la obtención de ventajas competitivas en los diferentes campos de acción.

Este impacto social y la rápida evolución de las tecnologías determinan también una nueva visión en el Derecho, que ha procurado responder a esta realidad como responsable por garantizar el uso adecuado de aquéllas en beneficio de la sociedad en general, y por la necesidad de satisfacer adecuadamente a las demandas de la sociedad que exige a los operadores del derecho, tanto judiciales como legislativos, respuestas más rápidas, más acertadas y en mayor cantidad, lo que hace que el estudio de herramientas informáticas aplicadas al mundo jurídico suponga una

acción muy importante, dentro del sector informático.

Por esta razón, este proyecto propone una investigación exploratoria del modelado y visualización de Big Data para documentación y datos jurídicos, basado en la tarea de realizar la recolección, análisis y crítica de los diferentes métodos propuestos en la literatura sobre el manejo de grandes volúmenes de datos en el proceso de razonamiento y respuesta a consultas, y su consolidación en un trabajo que relacione estas perspectivas diferentes, analizando las ventajas y desventajas, así como los retos de implementación que presentan.

Palabras clave: BigData, datos jurídicos, grandes volúmenes de datos, modelado, visualización.

CONTEXTO

La facultad de Ciencias para la Innovación y el Desarrollo de la Universidad Católica de Santiago del Estero impulsa la implementación de proyectos que apuntan a incentivar la investigación desde las cátedras promoviendo la interacción vertical y horizontal entre ellas, y posibilitando a los docentes obtener resultados que puedan ser aplicados en las aulas.

Por ello es que, a fines del año 2017, desde las asignaturas de Programación I, Estructuras de Datos y Análisis Numérico surge el proyecto “Big Data: Modelado y visualización de grandes volúmenes de datos jurídicos” con el objetivo de promover la investigación aplicada, la formación de recursos humanos, la innovación de los

contenidos de las cátedras y de las prácticas profesionales.

Este artículo presenta una línea de investigación que surge de ese proyecto, ante la necesidad de los operadores del derecho de disponer de instrumentos capaces de brindar el aprovechamiento de las oportunidades que genera el Big Data como herramienta predictiva y en la tarea de toma de decisiones.

1. INTRODUCCIÓN

La gran cantidad de información disponible y su heterogeneidad han sobrepasado la capacidad de las tecnologías actuales de gestión de datos. El tratamiento con grandes volúmenes de datos estructurados y no estructurados, a menudo referido como Big Data, es un tema de investigación de actualidad así como un importante desafío tecnológico.

Por tratarse de un fenómeno reciente, Big Data no se halla todavía suficientemente delimitado, aunque se puede encontrar varios puntos en común entre las definiciones planteadas por la comunidad académica, los técnicos y los profesionales especializados en el área. McKinsey [1] ofrece una definición “intencionalmente subjetiva” y dinámica a partir de las limitaciones tecnológicas en la actualidad: “Big Data refiere a un conjunto de datos cuyo tamaño está más allá de la capacidad que tienen los software de base de datos para capturar, almacenar, administrar y analizar”. De esta manera, a medida que la frontera tecnológica se va modificando, el tamaño del conjunto de datos que califica como Big Data también irá cambiando. Se habla de Big Data cuando el tamaño se vuelve el principal problema. Si bien el volumen es un punto en común entre los especialistas que definen Big Data, otros conceptos incorporan más dimensiones al análisis. Gartner [2] y UN Global Pulse [3] suelen caracterizar el tema a partir de los desafíos que Big Data representa no sólo a partir del volumen, sino también de la variedad de la información y la velocidad

de su tratamiento¹, incluso dándole mayor importancia a estas últimas. Con *variedad* se suele hacer referencia a la heterogeneidad de la representación e interpretación semántica de los datos, es decir que no están directamente listos para ser integrados a una aplicación; mientras que con *velocidad* se hace referencia tanto a la frecuencia de los datos como al tiempo de respuesta: la importancia reside en la velocidad del feedback, utilizando los datos para la toma de decisiones.

En sus primeros años, la informática jurídica fue, sustancialmente, informática documental, y abarca la creación, gestión y recuperación de datos en bancos de datos que contienen informaciones específicamente jurídicas, tales como leyes, doctrina y jurisprudencia, que fue recibida de buen grado por abogados y juristas.

Pero ante la existencia de un gran volumen de documentos e informaciones jurídicas, es necesario que los operadores del derecho dispongan de instrumentos capaces de brindar las oportunidades que genera el Big Data, como una herramienta predictiva de auxilio a sus tareas decisorias.

Frente a este desafío surgen las siguientes preguntas: ¿se encuentra el país, y en particular los poderes judiciales provinciales, en condiciones de aprovechar los beneficios que promete esta nueva fase de grandes volúmenes de datos, basados en el conocimiento?; si el análisis de grandes datos requiere de una base tecnológico-comunicacional: ¿están los poderes judiciales en condiciones de dar respuesta a esta exigencia?; ¿cómo impacta el escaso desarrollo local de infraestructura de alta complejidad para el desarrollo de Big Data?

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación del presente trabajo tiene como eje central el estudio y análisis de

¹ Volumen, Velocidad y Variedad, tres términos conocidos como las 3 Vs

las técnicas de visualización de grandes volúmenes de datos jurídicos.

Entre los supuestos que guían el proyecto se encuentran:

- En las últimas décadas, el sector de servicios de informática e información de la Argentina se posicionó entre los más dinámicos de la región [4], por lo que existen capacidades técnicas reales y potenciales que permiten aprovechar este avance.
- Asociados al fenómeno del Big Data existen tres tipos de problemas: *tecnológicos*, relacionados con el almacenamiento, la seguridad y el análisis de los volúmenes crecientes de datos; *comerciales*, relacionados con el valor añadido que se puede generar; y *sociales*, relacionados con la privacidad de la información personal [5].
- Los retos que genera Big Data están asociados a cambios de paradigma que implican una mayor importancia de la disponibilidad y el acceso a los datos, aceptación de ciertos niveles de imprecisión y desorden, y concentración en las correlaciones. [6]

Se trabajará en modelos de repositorios institucionales que se ajusten a los requerimientos de visualización de Big Data de carácter jurídico.

3. RESULTADOS ESPERADOS

Considerando que el mercado general de Big Data en Argentina muestra una demanda escasa y de baja complejidad, explicado en parte por la falta de conocimiento sobre el tema, problemas institucionales y limitaciones de infraestructura [7], el objetivo general del proyecto es realizar un estudio y análisis de técnicas de modelado y visualización de grandes volúmenes de datos jurídicos.

Este objetivo general se desarrollará mediante los siguientes objetivos específicos:

- Entender la problemática de la gestión del Big Data.
- Identificar las características más relevantes en la gestión de Big Data que deben guiar la elección de una solución para el ámbito jurídico.
- Analizar las principales herramientas de gestión de Big Data existentes en el mercado.
- Identificar los modelos más adecuados para el problema de los grandes volúmenes de datos jurídicos.
- Adquirir conocimientos específicos sobre el uso de Big Data para la toma de decisiones.
- Emplear herramientas de modelización de grandes volúmenes de datos y documentación del ámbito jurídico.

Para alcanzar el objetivo planteado, se llevarán a cabo las actividades siguientes:

- Buscar, recopilar y clasificar publicaciones, artículos y libros relacionados con la problemática planteada.
- Investigar los recursos técnicos existentes en el mercado, estudiando las metodologías de modelado, las plataformas y software para el tratamiento de Big Data y los diferentes métodos de visualización, estableciendo una comparación entre ellos.
- Identificar las características más relevantes de la gestión de Big Data que deben guiar la elección de una solución arquitectónica para datos jurídicos, seleccionando los modelos estadísticos o de aprendizaje automatizado más adecuados para el problema.
- Validar la solución escogida, evaluando el impacto de la utilización de Big Data en el ámbito jurídico.

Se espera que los resultados de esta investigación se incorporen a los contenidos de las cátedras relacionadas y al espacio curricular correspondiente.

Como resultados indirectos se espera la consolidación del grupo de investigación, la formación de nuevos investigadores y la motivación y entrenamiento en investigación de los estudiantes de grado.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo está conformado por tres docentes de la carrera de Ingeniería en Informática, con dedicación simple.

El grupo hace difusión y formación de recursos humanos desde las cátedras: Programación I, Estructura de Datos y Análisis Numérico.

Asimismo, se considera de gran interés la incorporación de becarios, para motivar a los alumnos de la carrera de Ingeniería en Informática a realizar su trabajo final de grado en el área de este proyecto.

5. BIBLIOGRAFIA

- 1] MCKINSEY, (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute.
- 2] GARTNER, (2011). Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data
- 3] UN GLOBAL PULSE (2012). Big Data for Development: Challenges & Opportunities. United Nations.
- 4] BARLETTA, PEREIRA, ROBERT & YOGUEL, (2013). "Argentina: Dinámica reciente del sector de software y servicios informáticos", Revista CEPAL N° 110
- 5] NUNAN, DI-DOMENICO, (2013). "Market research and the ethics of big data". International journal of market research, v. 55, n. 4, pp. 505-520. Disponible en: <http://dx.doi.org/10.2501/IJMR-2013-015>
- 6] MAYER-SCHÖNBERGER, CUKIER (2013). Big data. La revolución de los datos masivos. Madrid: Turner
- 7] MALVICINO & YOGUEL (2015). "Descubriendo Big Data en Argentina. Encuesta Digital 2014". AGRANDA 1era ed. 44va Jornadas de Informática (JAIIO)

DISEÑO Y CONSTRUCCIÓN DE PROCESOS DE EXPLOTACIÓN DE INFORMACIÓN PARA EL ÁREA DE CIENCIAS DE LA COMPUTACIÓN

H. Kuna, M. Rey, E. Zamudio, A. Canteros, A. Cantero, A. Rambo, E. Martini, G. Pautsch, C. Biale, S. Krujoski, F. Rauber

Departamento de Informática, Facultad de Ciencias Exactas Químicas y Naturales, Universidad Nacional de Misiones.

hdkuna@gmail.com

RESUMEN

El desarrollo de procesos de explotación de información aplicado a la recuperación de producciones científico-tecnológicas, exige resolver problemáticas específicas relacionadas con las entidades del dominio, tales como publicaciones, instituciones, y autores.

Continuando con la línea general de este grupo de investigación, este trabajo aborda dos problemáticas actuales para la adecuada gestión de un servicio de recuperación de información en general, y del área de las Ciencias de la Computación en particular. Estas problemáticas incluyen la desambiguación y la recomendación de entidades asociadas a las producciones científico-tecnológicas.

El abordaje propuesto de estas problemáticas se enfoca principalmente en la aplicación de técnicas como el Procesamiento de Lenguaje Natural, Aprendizaje de Máquina y Análisis de Redes Sociales.

Estas propuestas prevén la evaluación experimental en el contexto de un servicio de metabuscador de publicaciones científicas del área de las Ciencias de la Computación, desarrollado y mantenido por este grupo de investigación.

Mediante este trabajo se pretende contribuir en la mejora del desempeño de procesos actuales de explotación de información asociados a la recuperación de producciones científico-tecnológicas.

Palabras clave: producciones científico-tecnológicas, recuperación de información, desambiguación, recomendación

CONTEXTO

Esta línea de investigación articula el Programa de Investigación en Computación (PICom) de la Facultad de Ciencias Exactas Químicas y Naturales, Universidad Nacional de Misiones (FCEQyN/UNaM) con el grupo de investigación Soft Management of Internet and Learning (SMILe) de la Universidad de Castilla-La Mancha, España, y con el Departamento de Matemáticas de la Universidad de Sonora, México.

1 INTRODUCCIÓN

La recuperación de producciones científico-tecnológicas de relevancia a través de la web es uno de los desafíos actuales que involucra la actividad científica. El volumen de publicaciones disponibles, la variedad de fuentes y los datos que se generan a partir de la interacción entre publicaciones, autores y demás entidades involucradas, constituyen los ejes de este problema. En la actualidad, existen distintas alternativas que permiten a los usuarios acceder a tales contenidos, parte de ellas son las herramientas de búsqueda web y los repositorios digitales.

Este tipo de herramientas, se encuentran en auge en la actualidad. En particular, el área de las Ciencias de la Computación agrupa organizaciones que implementan servicios de búsqueda de algún tipo [1], [2] como son los repositorios de instituciones como ACM, IEEE, Elsevier, Springer, ArXiv, DBLP y DOAJ, entre otros; los motores de búsqueda de

Google Scholar y Microsoft Academics [3]; y los repositorios digitales de documentos científicos que son gestionados por instituciones educativas o científicas, y que contienen datos primarios de su producción científico-tecnológica [4].

En nuestro país, a partir de la creación del Sistema Nacional de Repositorios Digitales, se cuenta con una herramienta de búsqueda, que incluye más de 100.000 documentos de diferentes tipos y de todo el país [5].

Este contexto evidencia la necesidad de enfoques que asistan al proceso de recuperación de información a través de la implementación de funcionalidades destinadas a satisfacer las necesidades de los usuarios. Así, los procesos de explotación de información pueden servir de soporte para acciones como: integración de resultados, especificación del orden de relevancia de los resultados, clasificación automática de los materiales, visualización de redes de colaboración entre entidades (autores, equipos de investigación, instituciones) y la identificación unívoca de contenidos, entre otros.

Este problema ha sido abordado previamente por este grupo de investigación desde el área de Recuperación de Información (RI), particularmente dirigido a la búsqueda de producciones científicas en el área de Ciencias de la Computación.

Como consecuencia de trabajos previos, este grupo ha desarrollado un metabuscador de producciones científicas para el área de las Ciencias de la Computación. Este metabuscador ha experimentado varias transformaciones desde el momento de su concepción [9]–[12].

Entre los elementos o módulos que conforman el mencionado metabuscador, se incluyen: un algoritmo de ranking basado en las propiedades de las publicaciones recuperadas [9], un método de expansión de consultas basado en ontologías de un área disciplinar [10] y un conjunto de operaciones destinadas a la gestión de los datos de las entidades con las que opera el metabuscador [11], [12].

Estos avances han contribuido significativamente a mejorar la recuperación de publicaciones científicas en el área de las Ciencias de la Computación. Sin embargo, estos mismos avances han planteado nuevos desafíos en relación a la gestión interna del metabuscador, como así también a los servicios ofrecidos por éste.

Por una parte, se evidencia la necesidad de gestionar adecuadamente las entidades que almacena el metabuscador para asistir a los procesos de búsqueda, recuperación, y presentación de resultados. Esta gestión de las entidades, requiere implementar procesos de desambiguación que ayuden a determinar unívocamente aquellas entidades asociadas a las producciones científicas. Estas entidades incluyen autores, instituciones, y títulos, entre otros.

Por otra parte, es necesario asistir al usuario del metabuscador a partir de recomendaciones de producciones científicas que pudieran resultar relevantes para su investigación. Esta tarea requiere implementar procesos de recomendación que ayuden a identificar entidades que podrían ser relevantes para el investigador. En particular, la recomendación de autores requiere identificar, entre otros aspectos, las áreas de experiencia de dichos autores. Así es que resulta necesario también investigar sobre los temas asociados a los autores de las producciones científicas.

1.1 RECOMENDACIÓN MEDIANTE PERFILES DE EXPERTOS

En particular, cuando las entidades recomendadas son autores, el sistema de recuperación de información puede intentar determinar las áreas de conocimiento asociadas a dichos autores. Esta tarea se conoce como Perfilado de Expertos o Expert Profiling. La construcción de Perfiles de Expertos se encuentra estrechamente relacionada con las fuentes de datos disponibles para determinar la evidencia de experiencia de las personas. En el ámbito académico, pueden utilizarse varias fuentes de evidencia de experiencia, como los CV, páginas personales, y otros. Sin embargo, una de las principales fuentes de evidencia de

experiencia son las mismas producciones científico-tecnológicas de los autores.

El principal problema de utilizar producciones científico-tecnológicas como fuente primaria de experiencia en la construcción de perfiles de expertos, es que estas producciones se componen principalmente por texto, el cual debe procesarse de algún modo, para extraer algún tipo de información.

La generación automática de perfiles de expertos en el ámbito académico utilizando producciones científico-tecnológicas como principal fuente de experiencia, implica la necesidad de desarrollar y evaluar técnicas avanzadas de computación, principalmente pertenecientes al área del Procesamiento de Lenguaje Natural y del Aprendizaje de Máquina.

El desarrollo de estrategias para la generación automática de perfiles de expertos se presenta como una alternativa para ofrecer nuevas funcionalidades a los sistemas de recuperación de información en el ámbito académico. Asimismo, la generación de los perfiles de expertos en ámbitos académicos permite extender sus aplicaciones a problemas frecuentes del contexto, como ser la recomendación de expertos para evaluación de proyectos de investigación, becarios, asignación de subsidios, o conformación de paneles de expertos para el tratamiento de temas específicos o consultoría. Asimismo, y considerando los ámbitos de los primeros desafíos en la generación de perfiles de expertos relacionados con el ámbito empresarial, las contribuciones alcanzadas en un contexto incentivan su adaptación y aplicación en otros contextos como la selección de expertos para el tratamiento de temas específicos, selección de personal en el ámbito laboral, y selección de comités, entre otros.

1.2 DESAMBIGUACIÓN DE AUTORES

Cuando se publica una producción científica, se desea conocer al autor de la misma, y en algunos casos, si éste ha publicado otras obras. Con la masividad de información en internet, dicha búsqueda se vuelve más compleja. Una tarea que antes era llevada a cabo manualmente analizando datos descriptivos del

autor, y a veces sus textos, es casi imposible de realizar hoy en día.

La actividad para solucionar esta problemática es lo que se conoce como desambiguación de autores.

Algunos ejemplos en que se requiere desambiguar los autores, incluyen: publicaciones de autores con distinto nombre, o diferentes autores con un mismo nombre. Además, puede ocurrir que la información sobre el autor sea insuficiente o presente inconsistencias en su identificación. Por ejemplo: algunas editoriales no publican el primer nombre de los autores, o su información geográfica, títulos que poseen o áreas en las que son expertos, entre otros.

En un Sistema de Recuperación de Información (SRI), la identificación del autor es necesaria en la aplicación de las métricas de calidad, por ejemplo, en un algoritmo de ranking.

Los desafíos que presenta la desambiguación de autores han llevado al desarrollo de varios métodos. En [10], se presenta una taxonomía jerárquica caracterizando los distintos métodos, y haciendo referencia a aquellos que son más representativos. De acuerdo con esta taxonomía, los métodos se clasifican: según su aproximación, en agrupamiento de autores [11]–[14] y asignación de autores. Esta última a su vez se divide en clasificación [15] y clustering [16]. Alternativamente, estos métodos se pueden clasificar según la evidencia explorada en aquellos que utilizan datos presentes en las citas, los que utilizan información en la web [13], y los que extraen datos implícitos a partir de lo que esté disponible [14].

2 LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La presente línea de investigación propone como objetivo general, el desarrollo de procesos de explotación de información para su implementación en un sistema de recuperación de información de producciones científicas del área de las Ciencias de la Computación.

Este objetivo deriva en dos aspectos principales. El primer aspecto consiste en investigar sobre los procesos de recomendación automática de datos asociados a producciones científicas, como autores o publicaciones. El segundo aspecto consiste en investigar sobre los procesos de desambiguación de entidades correspondientes a producciones científicas, principalmente autores.

El aspecto de recomendación, será abordado utilizando datos internos del metabuscador, con el objetivo de analizar los perfiles de los autores ahí almacenados [8, 9].

Se prevé la evaluación de un conjunto de algoritmos de recomendación para ser aplicado sobre un conjunto de autores que podrían ser de interés para el usuario del metabuscador, a partir de la consulta que haya ingresado.

El aspecto de desambiguación, será abordado a partir de técnicas integradas de procesamiento de lenguaje natural, aprendizaje automático, y análisis de redes sociales, a partir de los datos extraídos de las producciones científicas.

Ambos aspectos de la línea de investigación general, prevé evaluaciones de tipo experimental en el contexto del metabuscador ya desarrollado.

3 RESULTADOS Y OBJETIVOS

El desarrollo de los procesos de explotación de información, que conforman el objetivo principal de esta línea de trabajo, incluye los siguientes objetivos particulares:

- Desarrollar estrategias que permitan la generación automática de perfiles de expertos en el ámbito de sistemas de recuperación de información académicos mediante la aplicación y desarrollo de técnicas avanzadas de Aprendizaje de Máquina y Procesamiento de Lenguaje Natural.
- Desarrollar métodos de recomendación de contenido para las entidades con las que opera el metabuscador, siguiendo la línea del sistema de recomendación de autores,

brindando resultados en forma paralela a la ejecución de las consultas del usuario.

- Desarrollar métodos para la gestión de los datos con los que opera el metabuscador, tal como el de desambiguación de entidades, detección de outliers, y la especificación de una taxonomía propia para las entidades almacenadas, entre otros.
- Analizar la factibilidad de implementación de métodos basados en sistemas inteligentes para la determinación de la relevancia de los resultados del metabuscador, considerando los datos de los artículos, autores, lugares de publicación y otras entidades involucradas.
- Finalizar la implementación de un método que genere perfiles de los usuarios del metabuscador y los considere para la mejora de la experiencia del usuario con la herramienta.
- Evaluar el desempeño del meta-buscador en relación con soluciones de recuperación de información que operen sobre contextos similares.

4 FORMACIÓN DE RECURSOS HUMANOS

Este proyecto es parte de las líneas de investigación en recuperación y explotación de información del Programa de Investigación en Computación (FCEQyN/UNaM), con once integrantes relacionados con las carreras de Ciencias de la Computación de la UNaM. En resumen, el grupo de investigación desarrolla tres tesis de grado articulando sus trabajos con becas de Estimulo a las Vocaciones Científicas del Consejo InterUniversitario Nacional (CIN), tres tesis de maestría, y un trabajo de investigación posdoctoral. Asimismo, la línea y el equipo de investigación se vinculan con el Grupo de Investigación SMILe de la Universidad de Castilla-La Mancha, España y con la línea de trabajo en aprendizaje automático del Departamento de Matemáticas de la Universidad de Sonora, México.

5 BIBLIOGRAFÍA

- [1] M. E. Falagas, E. I. Pitsouni, G. A. Malietzis, and G. Pappas, "Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses," *FASEB J.*, vol. 22, no. 2, pp. 338–342, Feb. 2008.
- [2] M. Ley, "The DBLP Computer Science Bibliography: Evolution, Research Issues, Perspectives," in *String Processing and Information Retrieval*, A. H. F. Laender and A. L. Oliveira, Eds. Springer Berlin Heidelberg, 2002, pp. 1–10.
- [3] J. L. Ortega and I. F. Aguillo, "Microsoft academic search and Google scholar citations: Comparative analysis of author profiles," *J. Assoc. Inf. Sci. Technol.*, vol. 65, no. 6, pp. 1149–1156, 2014.
- [4] J. Tang, "AMiner: Mining Deep Knowledge from Big Scholar Data," in *Proceedings of the 25th International Conference Companion on World Wide Web*, Republic and Canton of Geneva, Switzerland, 2016, pp. 373–373.
- [5] *Sistema Nacional de Repositorios Digitales, & Ministerio de Ciencia*. Portal de Datos. URL: <http://repositoriosdigitales.mincyt.gob.ar>
- [6] H. Kuna, E. Martini, and M. Rey, "Evolution of a Ranking Algorithm for Scientific Documents in the Computer Science Area," in *XX Argentine Congress of Computer Science Selected Papers*, La Plata, Buenos Aires, Argentina: EDULP, 2015, pp. 145–155.
- [7] H. Kuna, M. Rey, L. Podkowa, E. Martini, and L. Solonezen, "Expansión de Consultas Basada en Ontologías para un Sistema de Recuperación de Información," presented at the XVI Workshop de Investigadores en Ciencias de la Computación, 2014.
- [8] M. Rey *et al.*, "Propuesta de Esquemas de Perfiles para la Recuperación de Datos Científicos para un Sistema de Recuperación de Información del Área de Ciencias de la Computación," presented at the XXII Congreso Argentino de Ciencias de la Computación, San Luis, Argentina, 2016.
- [9] H. Kuna *et al.*, "An Entity Profile Schema for Data Integration in an Academic Metasearch Engine," in *Proceedings of the 2017 International Conference on Artificial Intelligence*, Las Vegas, USA, 2017, pp. 281–285.
- [10] A. A. Ferreira, M. A. Gonçalves, and A. H. F. Laender, "A Brief Survey of Automatic Methods for Author Name Disambiguation," *SIGMOD Rec*, vol. 41, no. 2, pp. 15–26, Aug. 2012.
- [11] Y. Liu, W. Li, Z. Huang, and Q. Fang, "A fast method based on multiple clustering for name disambiguation in bibliographic citations," *J. Assoc. Inf. Sci. Technol.*, vol. 66, no. 3, pp. 634–644, Mar. 2015.
- [12] M. Khabsa, P. Treeratpituk, and C. L. Giles, "Large scale author name disambiguation in digital libraries," in *2014 IEEE International Conference on Big Data (Big Data)*, 2014, pp. 41–42.
- [13] Y. Qian, Q. Zheng, T. Sakai, J. Ye, and J. Liu, "Dynamic Author Name Disambiguation for Growing Digital Libraries," *Inf Retr*, vol. 18, no. 5, pp. 379–412, Oct. 2015.
- [14] A. F. Santana, M. A. Gonçalves, A. H. F. Laender, and A. A. Ferreira, "Incremental Author Name Disambiguation by Exploiting Domain-specific Heuristics," *J Assoc Inf Sci Technol*, vol. 68, no. 4, pp. 931–945, Apr. 2017.
- [15] A. A. Ferreira, A. Veloso, M. A. Gonçalves, and A. H. F. Laender, "Self-training author name disambiguation for information scarce scenarios," *J. Assoc. Inf. Sci. Technol.*, vol. 65, no. 6, pp. 1257–1278, Jun. 2014.
- [16] J. Tang, A. C. M. Fong, B. Wang, and J. Zhang, "A Unified Probabilistic Framework for Name Disambiguation in Digital Library," *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 6, pp. 975–987, Jun. 2012.

Hacia la optimización del uso de datos abiertos en el ámbito público

Carrizo Claudio*, Pablo Vaca*, Salgado Carlos+, Sánchez Alberto+, Peralta Mario+

*Facultad Regional San Francisco – Facultad Regional Córdoba

Universidad Tecnológica Nacional

Av. de la Universidad 501 - San Francisco - Córdoba - Tel. 03564-421147

{jcarrizo77, vacapablo72}@gmail.com

+ Departamento de Informática Facultad de Ciencias Físico-Matemáticas y

Naturales Universidad Nacional de San Luis

Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina

e-mail: {csalgado, alfanego, mperalta}@unsl.edu.ar

RESUMEN

En la actualidad, existe una fuerte iniciativa por parte de los gobiernos en poner los datos públicos a disposición de los ciudadanos, con el fin de que los mismos puedan aportar un valor a la información, en pos de mejorar la calidad de vida de los habitantes, impulsando lo que se denomina “Ciudades Inteligentes”.

Pero desde la experiencia de los ciudadanos que consumen estos datos, existen dificultades al realizar procesos ETL (Extraction, Transformation and Load, en inglés) sobre todo, en lo que respecta a normalización, formatos de presentación e interpretación de los datos.

El objetivo de esta investigación consiste en proponer un modelo y algunos formatos de datos que sean apropiados al dominio del problema. El modelo de datos debe permitir caracterizar y normalizar los datos, mientras que los formatos recomendables junto con la exposición de una metadata asociada, deben permitir optimizar la presentación e interpretación de los datos abiertos u OD (Open Data, en inglés).

Este trabajo está enfocado en los portales web de datos abiertos del ámbito público, en pos de que estos puedan brindar datos que sean óptimos para ser utilizados en

iniciativas innovadoras, que puedan contribuir a mejorar la calidad de vida de los ciudadanos.

Palabras clave:

Datos Abiertos, Ciudades Inteligentes, Modelo de datos, Representación de datos, Metadata.

CONTEXTO

La presente línea de investigación se encuentra inmersa dentro del Proyecto de Investigación y Desarrollo Inter-facultad “Estado del Arte en Ciencia de Datos y Big Data”, el cual se encuentra homologado y financiado por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional, bajo el código IFN4567 y según la Disposición SCTyP N° 468/16. El periodo de ejecución de dicho proyecto es desde el 1 Enero de 2017 hasta el 31 de Diciembre de 2018, y el mismo está incluido en el Programa I&D + i de Tecnología de las Organizaciones de la Universidad Tecnológica Nacional.

1. INTRODUCCIÓN

Con el advenimiento de las nuevas tecnologías de la información y comunicaciones (TICs) [1] [2], existe una tendencia a nivel mundial acerca de la apertura de datos desde distintas instituciones, en especial las que pertenecen al estado nacional en todos sus

niveles. En los últimos años, los gobiernos a nivel de provincias/estados y países del mundo, han avanzado en la publicación de OD [3] [4] [7], no solo como un medio para generar transparencia, sino también para alentar su uso en iniciativas innovadoras que busquen mejorar la calidad de vida de los habitantes en las ciudades, para de esta poder contribuir en la búsqueda de lo que se denomina “Ciudades Inteligentes” [5] [6]. Una ciudad es considerada inteligente cuando aplica las TICs con el fin de dotar de una infraestructura que mejore la calidad de vida de sus habitantes y sea sostenible en el tiempo.

Según un trabajo publicado en la Academia de Estudios Económicos de Bucarest [7], OD es el concepto o la idea de aquellos que piensan que los datos deben poder ser accedidos por cualquier usuario y republicados todas las veces que se quiera, sin restricción de ningún tipo. Según Maximiliano Bron en el libro “Open Data, Miradas y perspectivas de los datos abiertos” [3], hablar de OD es mucho más que una creencia o un concepto, es una práctica o una filosofía que establece que los datos deben estar disponibles sin restricciones de acceso a ellos y que es algo similar a lo que ocurre con el software libre o de código abierto.

Las iniciativas OD son movilizadas por varios motivos, según un trabajo presentado por Calvin Chan en la Conferencia Internacional sobre Ciencias de Sistemas en el 2013 [4], hay dos causales principales para abrir la información, uno de ellos es la democracia y la libertad de la información, buscando ser más abiertos y transparentes, y el otro es económico, buscando que iniciativas privadas agreguen valor a la información.

Uno de los puntos que nos motivaron para trabajar en esta temática, fue el análisis de la forma con que los distintos actores de la sociedad se apropian de los conocimientos. Esto se sustenta con la experiencia que tienen los distintos integrantes del grupo y la

información recabada con los distintos actores de la empresa del medio con los que se interactúa por razones laborales, capacitaciones y estudio de posgrado. Otro de los aspectos que motivan el desarrollo de esta línea de investigación, está relacionado con las dificultades constantes que encuentran los alumnos al hacer uso de OD en el desarrollo de trabajos prácticos en la Cátedra de Big Data [8]: Arquitecturas y Estrategias de Análisis de Datos Masivos de la carrera de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Regional Córdoba. El desarrollo de estos trabajos consiste en buscar fuentes de datos abiertas en portales web en el ámbito público, en donde a través de procesos de ETL [9], se puedan obtener OD que luego serán utilizados para generar información de valor, a través de la aplicación de técnicas de Big Data o Minería de datos, en pos de lograr impacto positivo en las sociedades, mejorando de esta manera la calidad de vida de los ciudadanos. Del análisis de las dificultades encontradas por los alumnos en el uso de OD, se evidenciaron problemas de normalización, formatos de presentación e interpretación del significado de los datos abiertos.

Una posible solución al problema detectado anteriormente consiste en:

- Utilizar modelos que permitan caracterizar y normalizar OD.
- Identificar formatos de presentación según el origen y contexto de los datos
- Exponer una metadata asociada que permita interpretar el significado de los OD.

El objetivo de este trabajo consiste en realizar una discusión acerca del uso de modelos, formatos recomendables de presentación y exponer una metadata asociada para datos abiertos en el ámbito público.

Para lograr esto será necesario:

- Identificar y describir modelos que permitan caracterizar y normalizar los datos abiertos
- Analizar y evaluar distintos portales web que brinden datos abiertos públicos para identificar formatos de presentación recomendables y exposición de metadata.

Con este trabajo se propone realizar un aporte hacia los portales web de datos abiertos en el ámbito público, en pos de que los ciudadanos puedan hacer un mejor uso de los OD y de esta manera, puedan aportar valor a la información a través de sus iniciativas, contribuyendo a mejorar la calidad de vida de los habitantes en las ciudades.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los principales ejes de trabajo de esta línea de investigación se detallan a continuación:

- Identificar y describir modelos que permitan caracterizar y normalizar datos
- Analizar y evaluar portales de datos abiertos del ámbito público a nivel nacional e internacional
- Identificar formatos de presentación de datos según el origen y contexto de los datos
- Identificar formatos de exposición de metadata que permitan interpretar el significado de los OD
- Realizar una discusión acerca del uso de modelos y el formato de presentación de OD.

3. RESULTADOS OBTENIDOS/ESPERADOS

Hasta el momento, se ha logrado identificar y describir los modelos más utilizados, entre ellos, Modelo Relacional [10], Clave-Valor, Documental [11] [12]), los cuales permitirán caracterizar y normalizar datos en pos de evitar inconsistencias y redundancias e incrementar la velocidad de respuesta sobre las consultas u operaciones que se realizan sobre ellos.

Para lograr una mejor interpretación del significado de los datos, será necesario considerar el uso de una metadata [13]. Se puede tomar como referencia la plataforma Twitter [14] [15], dicha red social permite, mediante la tecnología REST (Representational State Transfer, en inglés), utilizar la información almacenada mediante el formato JSON [16], el cual no solo contiene los valores de los campos, sino también el nombre de los mismos. Otra alternativa presentada en el trabajo “Open Metadata Formats: Efficient XML-Based Communication for Heterogeneous Distributed Systems” [17], consiste en utilizar documentos XML como metadata, de tal manera que describan los datos expuestos, tal como el nombre, tipo de dato, extensión, significado o dominio. De la revisión del trabajo de Huamin Wang y Zhiwei Ye [18], se puede observar la utilidad de contar con una metadata. En dicho trabajo, se plantea un servicio de ETL basado en metadata, la misma es utilizada para especificar las estructuras de datos y además las fuentes de datos, reglas de procesamiento o relaciones entre ellos. De esta idea, surge la importancia de que al publicar datos abiertos, estos cuenten con una metadata asociada, lo cual facilita la utilización de los mismos y asegura su correcto entendimiento.

Lo anterior se ve reflejado en el estudio y análisis de distintos portales a nivel nacional e internacional de datos en ámbitos públicos. Comparando, por ejemplo, el portal de Aragón Open Data [19] y el de CABA [20], se observa que en ambos se utilizan los formatos estándares y los exponen con una metadata mínima. De la comparación entre ambos, se puede decir que el primero contiene repositorios más formateados, ya que todos los conjuntos de datos están en los formatos XML y JSON. En el segundo, no todos los conjuntos de datos son presentados con los mismos formatos, lo que agrega un nivel de dificultad extra al momento de utilizarlos.

Del estudio realizado sobre distintos portales se ha podido determinar que:

- En ambientes con información estructurada, se encontró cierta dificultad al momento de entender el significado de la información de los campos. Este tipo de formato debe ser mejorado agregando información sobre los datos, por ejemplo a través de la metadata de los mismos, utilizando por ejemplo un formato XML o similar que permita en el futuro agregar nuevas columnas de datos a los existentes.
- En ambientes con información no estructurada, como por ejemplo las redes sociales, o donde la estructura de la información cambia regularmente, es más adecuado un formato JSON, donde la metadata está incluida en cada registro, esto permitiría agregar o remover columnas sin afectar la información existente.

En resumen de lo antes expuesto, podemos concluir que los portales web de datos abiertos del ámbito público deberían considerar el uso de modelos que permitan caracterizar y normalizar datos, como así también permita sugerir formatos de presentación junto con la exposición de una metadata asociada en pos de que los ciudadanos puedan hacer un buen uso de los datos públicos y a través de sus iniciativas innovadoras puedan aportar información de valor que permita mejorar la calidad de vida de los habitantes en las ciudades.

4. FORMACIÓN DE RECURSOS HUMANOS

Esta línea de investigación se está llevando a cabo en forma conjunta entre la Universidad Nacional de San Luis (a través de la Facultad de Ciencias Físico-Matemáticas y Naturales) y la Universidad Tecnológica Nacional (a través de sus Facultades Regionales de Córdoba y San Francisco).

El equipo de trabajo está compuesto por 3

docentes investigadores categorizados a nivel nacional, 2 Tesistas de Posgrado y 2 becarios de grado que se encuentran cursando actualmente la carrera de Ingeniería en Sistemas de Información.

5. BIBLIOGRAFÍA

- [1] TIC - Tecnologías de la Información - <http://aprendeonline.udea.edu.co/lms/investigacion/mod/page/view.php?id=3118>.
- [2] Cortagerena Alicia, Freijedo Cludio – Tecnologías de la Información y la Comunicación – Editorial Prentice-Hall
- [3] Maximiliano Bron - Universidad Nacional de La Rioja - 2015 - Open Data, Miradas y perspectivas de los datos abiertos - Obra colectiva y colaborativa. 1era Edición. ISBN: 978-987-1999-09-5.
- [4] Calvin M.L. Chan - SIM University - From Open Data to Open Innovation Strategies: Creating e-Services Using Open Government Data. 2013 46th Hawaii International Conference on System Sciences.
- [5] Ciudad Inteligente - http://www.endesaeduca.com/Endesa_educarecursos-interactivos/smart-city
- [6] Ciudad Inteligente - <http://cintel.org.co/innovacion/ciudades-inteligentes>.
- [7] - Lorena BĂŢĂGAN - Academic Economics Studies - Bucarest - Rumania - Open Data for Smart Cities - lorena.batagan@ie.ase.ro.
- [8] Cátedra de Big Data - Universidad Tecnológica Nacional - Facultad Regional Córdoba. Profesor Titular: Ing. Calixto Maldonado. JTP: Ing. Franco Mana y Ayudante: Ing. Pablo Vaca.
- [9] Definición de ETL - <http://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/312584/procesos-etl-definicion-caracteristicas-beneficios-y-retos>.
- [10] Ramez Elmasri, Shamkant Navathe - Fundamentos de bases de datos - Quinta Edición - Pearson / Addison Wesley.

- [11] Neal Leavitt - Will NoSQL Databases Live Up to Their Promise?. TECHNOLOGY NEWS.
- [12] Karamjit Kaur - Rinkle Rani - Modeling and Querying Data in NoSQL Databases. Computer Sci. and Engg. Deptt. Thapar University.
- [13] José A. Senso, Antonio de la Rosa Piñero. El concepto de metadato. Algo más que descripción de recursos electrónicos.
- [14] Twitter para desarrolladores - <https://dev.twitter.com>.
- [15] Twitter - Red Social – <https://twitter.com>.
- [16] JSON Org. <http://json.org/example.html>.
- [17] Patrick Widener, Karsten Schwan y Greg Eisenhauer - College of Computing Georgia Institute of Technology - Atlanta, Georgia 30332-0280 - Open Metadata Formats: Efficient XML-Based Communication for Heterogeneous Distributed Systems.
- [18] Huamin Wang y Zhiwei Ye - International School of Software Wuhan University y School of Computer Science Hubei University of Technology - Wuhan, P.R. China - An ETL Services Framework Based on Metadata.
- [19] Aragón Open Data – <http://opendata.aragon.es>.
- [20] Portal de Datos – CABA - <https://data.buenosaires.gob.ar/>.

Técnicas de unificación de datos para la visualización de grandes volúmenes de datos

Lilia Palomo, Norma Lesca y Laura Sánchez Piccardi

Facultad de Matemática Aplicada - Universidad Católica de Santiago del Estero
lilia.palomo@ucse.edu.ar – norma.lesca@gmail.com – lsanchezpiccardi@gmail.com

RESUMEN

En la actualidad, la incorporación de las Tecnologías de la Información y la Comunicación (TIC) en el ámbito de empresas y organizaciones como un elemento clave para mejorar su competitividad e impulsar su crecimiento económico, ha producido un crecimiento considerable en el volumen de datos generados por diferentes sistemas y actividades, lo que hace dificultoso mantenerse al día con los resultados y genera la necesidad de modificar, optimizar y desarrollar métodos y modelos de almacenamiento y tratamiento de datos que suplan las falencias que presentan las bases de datos y los sistemas de gestión de datos tradicionales.

Aunque, se pueda contar con grandes repositorios de datos flexibles y multi-estructurados, que están disponibles a costos bajos o gratuitos y listos para ser explotados por procesos informacionales, la gestión de múltiples tipos de datos (incluidos los datos estructurados, semi-estructurados y no estructurados) se vuelve compleja, por la variedad o existencia de diferentes tipos y fuentes de datos. A lo que se suma, la necesidad de las organizaciones de integrar y analizar datos en un complejo abanico de fuentes de información tradicional y no tradicional, que son producidos tanto internamente como por fuera de la empresa.

El proceso de entrega (*delivery*) de datos para la toma de decisiones, continúa desempeñando un papel de innovación y, por lo general, se confunde con extracciones manuales de datos, repeticiones costosas de procesos, informes propensos a errores y una integridad de datos no mensurable. Pero en realidad facilita la

comprensión de los datos al transformarlos en información útil y ayuda a las organizaciones, a responder preguntas esenciales para la toma de decisiones que le permitan obtener ventajas competitivas y mejorar su posición en el mercado.

En este orden de ideas, cabe agregar como bien sostienen Jerry Held, Michael Stonebraker, Thomas H. Davenport, Ihab Ilyas, Michael L. Brodie, Andy Palmer, y James Markarian, 2016, que “la unificación de datos es una estrategia emergente, que cataloga el conjunto de datos, combina los datos de toda la empresa y los publica, para facilitar su consumo”. Es decir, se trata de tener el manejo y control de la información, a fin de tener asegurada una vista única de los datos, que provienen de fuentes funcionalmente distintas (bases corporativas, bases propias, sistemas externos, etc.), ya que los usuarios finales no tienen la necesidad de aprender a utilizar diferentes sistemas de acceso y manipulación de los datos. El uso de la unificación de datos como una estrategia *frontend* puede acelerar el suministro de datos altamente organizados en sistemas, como ETL (Extract, Transform and Load) y MDM (Master Data Management) y lagos de datos (data lake), aumentando el valor de estos sistemas y los conocimientos que permiten. [1] De lo expuesto precedentemente, se desprende el propósito del presente proyecto como una investigación exploratoria de las estrategias del proceso de unificación y delivery de grandes volúmenes de datos organizacionales.

Palabras clave: Unificación de datos, BigData, Integración de datos, Armonización de datos, Entrega de datos unificados.

CONTEXTO

El presente artículo presenta una línea de estudio, ante la necesidad de las organizaciones modernas de contar con un buen proceso de unificación de datos que les permita obtener el mayor provecho de ellos para mejorar la toma de decisiones.

Integra una línea del trabajo de investigación de cátedra, "Aproximación teórica de las estrategias de delivery de datos unificados del ámbito organizacional", que promueve la interacción vertical y horizontal, a partir de las asignaturas de Programación I, Estructuras de Datos, Análisis Numérico, Sistemas de Información e Ingeniería de Software, en el marco de la carrera de Ingeniería en Informática de la facultad de Ciencias para la Innovación y el Desarrollo perteneciente a la Universidad Católica de Santiago del Estero.

Posibilita a los docentes obtener resultados que puedan ser aplicados en las aulas con el objetivo de promover la innovación de los contenidos de las cátedras y de las prácticas profesionales.

1. INTRODUCCIÓN

En la era de Big Data todas las organizaciones, independientemente de su tamaño o industria, tienen dificultades para gestionar de manera efectiva grandes volúmenes de datos internos y externos. Cuando los datos se encuentran en diferentes bases de datos y otros repositorios, las empresas se ven obligadas a establecer múltiples integraciones e interfaces de usuario para construir una visión holística del rendimiento del negocio. El proceso de extracción e integración de datos se vuelve altamente complejo, y hasta puede tornarse completamente inmanejable debido a un ecosistema demasiado complicado, de herramientas de Business Intelligence (BI) superpuestas y sistemas de Planificación de Recursos Empresariales (ERP) dispares [2].

En consecuencia, las organizaciones no pueden construir una base sólida para la

estrategia de datos a largo plazo sin arriesgarse a una reducción de la alineación entre TI y los objetivos de la línea de negocio. Los riesgos operativos y analíticos incluyen [3]:

- Rendimiento operacional reducido.
- Implementaciones de TI reactivas y costosas.
- Falta de nuevos conocimientos.
- Informes y análisis caóticos
- Baja flexibilidad y tiempo de reacción a los cambios en el negocio.
- Problemas de calidad de datos no resueltos.
- Incapacidad para satisfacer las necesidades del cliente.
- Datos incorrectos y decisiones costosas.

Actualmente, la visión de casi todas las grandes organizaciones es maximizar el uso de sus activos de información para generar una ventaja competitiva. Sin embargo, solo unas pocas organizaciones modernas, realmente pueden sacar el máximo provecho de sus datos. A lo que se suma una tensión constructiva, ya que los consumidores de datos exigen un autoservicio instantáneo y siempre activo.

Hasta ahora, gran parte del enfoque, se ha relacionado con el uso de tecnologías de almacenamiento y herramientas analíticas para lograr este objetivo. Sin embargo, una pieza del rompecabezas que frecuentemente se pasa por alto, es el aprovechamiento de nuevos métodos para administrar los datos que conectan los sistemas de almacenamiento con los usos posteriores, como el análisis, ya que sin datos completos y limpios, el análisis se vuelve incompleto, inexacto e incluso engañoso [4].

En el ámbito del BI, el sofisticado trabajo de diseño de datos posibilita un análisis profundo. Por otro lado desde el entorno del workflow del proceso del Big Data, el almacenamiento de datos se hace más barato y más escalable.

Para llegar a una buena visualización, es

necesario entregar datos desde el origen hasta el punto de consumo. En el medio de este contexto, los procesos anticuados (tradicionales) pueden estorbar, por cuanto los datos se extraen manualmente y la falta de conexión puede producir errores. Además, las restricciones de ancho de banda retrasan la entrega del resultado de la consulta y reducen la puntualidad. Las evaluaciones cualitativas se transforman en hechos cuantitativos.

Entonces, el proceso semi-elaborado también puede llevar a la comprensión, pero hay que hacer una limpieza de los datos que tienen una confianza más débil, tarea que involucra mucho esfuerzo mal dirigido y ad-hoc.

Un buen Business Intelligence, puede ser una buena idea e idealmente conduce a un buen resultado. Pero cuando la estrategia se basa en las opciones de almacenamiento de datos, el flujo de entrada está restringido, la visibilidad se reduce y, en el peor de los casos, los resultados del negocio están predeterminados.

El solo hecho de que se hayan incorporado los datos no significa que se puedan operar o interpretar por un analista: los datos se encuentran en silos funcionales en diferentes ubicaciones de almacenamiento, o en silos lógicos dentro de un lago de datos no tan fácilmente accesible. La realidad es que los datos en un almacén de datos no pueden abrirse paso fácilmente, completamente curados e integrados a los consumidores. Para cuando se vuelve utilizable, a menudo se ha vuelto obsoleto [5].

En medio de esta situación, se encuentra como solución, el delivery, que sería la entrega de datos conectados o interrelacionados. Es hora de unificar los datos para que los usuarios comerciales puedan obtener el mayor provecho de los mismos.

Según señala, Michael Collins (2017), la tecnología de "unificación de datos", aprovecha las técnicas de aprendizaje automático y es una reinención de las capacidades tradicionales de gestión de datos, como las que se encuentran en los procesos de MDM y ETL, para cumplir los requisitos de la

era de Big Data [6].

Las técnicas tradicionales de gestión de datos son adecuadas cuando los conjuntos de datos son estáticos y relativamente pocos, pero fracasan en entornos de gran volumen y complejidad. Esto se debe en gran medida a sus enfoques descendentes y basados en reglas, que a menudo requieren un esfuerzo manual significativo para construir y mantener.

La tecnología de unificación de datos invierte este modelo, centrándose en la conexión y el dominio de conjuntos de datos mediante el uso de aprendizaje automático guiado por humanos, que aprovecha las señales en los datos para determinar cómo debe integrarse. El uso de la automatización guiada por la inteligencia humana para integrar y dominar los conjuntos de datos genera beneficios sustanciales en cuanto a velocidad, escala y flexibilidad del modelo de datos, al tiempo que garantiza los más altos niveles de precisión y confianza en los resultados. En su nivel más fundamental, la unificación de datos trae la promesa del aprendizaje automático a la preparación de conjuntos de datos a escala.

Sobre la base de lo expuesto, la unificación de datos beneficia a las organizaciones que buscan maximizar la entrega de valor derivado de los datos. Específicamente, la unificación de datos:

- Reduce los esfuerzos ad-hoc y únicos de integración de datos.
- Acelera la velocidad para obtener una idea.
- Mejora la confianza del usuario en los datos (al establecer una fuente única de verdad).
- Introduce flexibilidad en los workflow de datos para adaptarse al cambio.

La unificación de datos establece un modelo organizacional escalable con el crecimiento de los datos, a diferencia de "la vieja manera", una alternativa manual ad-hoc, que no lo hace.

Con la unificación de datos las personas dentro de la organización tendrán más posibilidades

de integrar datos en sus proyectos, compartir resultados y mantener los datos actualizados. Este ciclo de datos disponibles es clave para la innovación e intercambio de conocimientos.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación del presente trabajo tiene como eje central el estudio y análisis de las técnicas de unificación de datos para la entrega y visualización de grandes volúmenes de datos.

Se pueden mencionar los siguientes supuestos que dan estructura a la temática o campo de estudio del proyecto:

- Existen capacidades técnicas reales y potenciales que permiten aprovechar el avance del sector de servicios de informática e información, por cuanto, en las últimas décadas, la Argentina se posicionó entre los más dinámicos de la región [7].
- El mercado general de Big Data en Argentina muestra una demanda escasa y de baja complejidad, explicado en parte por la falta de conocimiento sobre el tema, problemas institucionales y limitaciones de infraestructura [8].
- Existen tres tipos de problemas asociados al Big Data: los tecnológicos, que se relacionan con el almacenamiento, la seguridad y el análisis de los volúmenes crecientes de datos; los comerciales, que se corresponden al valor añadido generado; y los sociales, relacionados con la privacidad de la información personal [9].
- Desde un punto de vista académico, Big Data genera estos retos que se vinculan a su vez a tres cambios de paradigma: mayor importancia de la disponibilidad y acceso de los datos; aceptación de niveles de imprecisión y desorden en los datos; centrarse más en las correlaciones, en vez de buscar constantemente la causalidad [10].

Se trabajará en la tipificación de técnicas de unificación de grandes volúmenes de datos organizacionales para posibilitar su entrega y visualización.

3. RESULTADOS ESPERADOS

El objetivo de esta línea de investigación plantea realizar un estudio y análisis del proceso de unificación de datos como estrategia para la recopilación, integración, preparación y entrega de los datos organizacionales.

Para lograr ese objetivo, este trabajo se centrará en los objetivos específicos:

- Investigar los diferentes enfoques técnicos disponibles actualmente para lograr el estado final deseado de conjuntos de datos limpios, precisos y consolidados
- Identificar las limitaciones existentes en las soluciones tradicionales para unificar grandes volúmenes y variedades de datos.
- Desarrollar una aproximación teórica para la recopilación, integración y preparación de datos organizacionales que posibilite una entrega eficiente de los mismos.

Las actividades que se llevarán a cabo son las siguientes:

- Estudio del proceso de unificación de datos.
- Análisis de las metodologías de preparación de datos variables.
- Análisis y estudio de herramientas de software y hardware.
- Comparación de los recursos existentes en el mercado.
- Selección de estrategias para la combinación de unificación y preparación de grandes volúmenes de datos.
- Identificación de los métodos de entrega eficiente de datos.

Se espera que los resultados de esta investigación se incorporen a los contenidos

de las cátedras relacionadas y al espacio curricular correspondiente.

Los resultados esperados respecto a la formación de recursos humanos son hasta el momento la consolidación del grupo de investigación, la formación de nuevos investigadores y la motivación y entrenamiento en investigación de los estudiantes de grado.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo está conformado por tres docentes de la carrera de Ingeniería en Informática, 2 dos con dedicación simple y una semiexclusiva.

El grupo hace difusión y formación de recursos humanos desde las asignaturas: Programación I, Estructura de Datos y Análisis Numérico, Sistemas de Información e Ingeniería de Software.

Asimismo, se considera de gran interés la incorporación de becarios, para motivar a los alumnos de la carrera de Ingeniería en Informática a realizar su trabajo final de grado en el área de este proyecto.

5. BIBLIOGRAFIA

- 1] Jerry Held, Michael Stonebraker, Thomas H. Davenport, Ihab Ilyas, Michael L. Brodie, Andy Palmer, and James Markarian, 2016. "Getting Data Right Volume and Variety". Copyright © 2016 Tamr, Inc.. O'Reilly Media, Inc. Disponible en: <https://www.tamr.com/landing-pages/getting-data-right>
- 2] 1010DATA, 2016. "Data Unification". Copyright © 2016 1010data Inc. Disponible en: https://www.1010data.com/media/1316/1010data_whitepaper_data_unification.pdf
- 3] Daily Data News. "What is Data Unification?", 2017. Disponible en: <https://dailydatanews.com/2017/12/04/data-unification>
- 4] Andy Oram, 2017. "Agile Data Mastering". Copyright © 2018 O'Reilly Media. Disponible en: <https://www.tamr.com/landing-pages/agile-data-mastering-report-3>
- 5] Toph Whitmore, 2017. "Connected Data Delivery: Combining Data Unification and Data Preparation". Copyright © 2017 Blue Hill Research. Disponible en: <https://www.tamr.com/whitepaper/connected-data-delivery-combining-data-unification-data-preparation/>
- 6] Michael Collins, 2017. "Data Unification: A New Path to Digital Transformation". Disponible en: <https://www.enterprisetech.com/2017/08/09/data-unification-new-path-digital-transformation/>
- 7] BARLETTA, PEREIRA, ROBERT & YOGUEL, 2013. "Argentina: Dinámica reciente del sector de software y servicios informáticos", Revista CEPAL N° 110.
- 8] MALVICINO & YOGUEL, 2015. "Descubriendo Big Data en Argentina. Encuesta Digital 2014". AGRANDA 1era ed. 44va Jornadas de Informática (JAIIO).
- 9] NUNAN, DI-DOMENICO, 2013. "Market research and the ethics of big data". International journal of market research, v. 55, n. 4, pp. 505-520. Disponible en: <http://dx.doi.org/10.2501/IJMR-2013-015>
- 10] MAYER-SCHÖNBERGER, CUKIER 2013. "Big data. La revolución de los datos masivos". Madrid: Turner.

Bases de Datos y Aplicación de Técnicas de Avanzadas

Edilma Olinda Gagliardi, Maria Gisela Dorzán, Maria Teresa Taranilla,

Pablo Rafael Palmero y Carlos Andrés Casanova

Departamento de Informática
Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis, Argentina
{oli, mgdorzan, tarani, prpalmero, cacasanova}@unsl.edu.ar

Gregorio Hernández Peñalver

Departamento de Matemática Aplicada
Facultad de Informática
Universidad Politécnica de Madrid, España
gregorio@fi.upm.es

RESUMEN

En el marco del proyecto *Tecnologías Avanzadas de Bases de Datos*, la línea de investigación *Bases de Datos Espaciales y Espacio Temporales* está orientada a vincular las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas. El objetivo general consiste en utilizar técnicas, métodos y herramientas para investigación de base y en la resolución de problemas de optimización en diversos dominios de aplicación.

Palabras clave: Bases de Datos, Geometría Computacional, Metaheurísticas, Bases de Datos Espaciales y Espacio Temporales.

CONTEXTO

El proyecto *Tecnologías Avanzadas de Bases de Datos* desarrolla actividades vinculadas al tratamiento de objetos de diversos tipos, estructurados y no estructurados que son de utilidad en diversos campos de aplicación, tales como computación móvil, sistemas de información geográfica, computación gráfica, robótica, diseño asistido por computadora, motores de búsqueda en internet, entre otras.

En el proyecto coexisten tres líneas de investigación, orientadas al desarrollo de nuevos modelos para administrar y recuperar

información almacenada en repositorios de datos no estructurados, donde los escenarios de exploración requieren modelos tales como las bases de datos de texto, bases de datos espaciales, espacio temporales, bases de datos de imágenes, bases de datos de sonidos, espacios métricos, entre otros.

Algunas aplicaciones de bases de datos espaciales y espacio temporales necesitan guardar y consultar información actual e histórica acerca de posiciones referenciadas espacialmente y cambios de forma que tuvieron los objetos de estudio en diferentes escenarios a lo largo del tiempo, por lo que son necesarios estos modelos de bases de datos.

En este contexto, resulta apropiado vincular las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas, a los efectos de utilizar métodos y herramientas para la resolución de problemas en diversos dominios de aplicación y para la resolución de problemas de complejidad NP orientados a optimización.

Asimismo, se hace necesaria la incorporación de métodos científicos, procesos y sistemas de descubrimiento de información oculta en grandes cantidades de datos estructurados y no estructurados, que permitan obtener información valiosa mediante análisis

avanzados que traten diferentes tipos de datos, a los efectos de contribuir en la toma de decisiones en un sistema.

Por otra parte, con el avance de las tecnologías, situaciones del mundo real ameritan el uso de las mismas, en pos de una mejora sustancial en cuestiones de calidad, gestión, economía, etc. A nivel mundial, las Tecnologías de la Información y la Comunicación (TIC) continúan introduciendo cambios políticos, económicos, sociales y culturales, entre muchos otros, convirtiéndose en procesos clave para cualquiera que quiera accionar sobre el presente y proyectar hacia el futuro.

Por lo expuesto, en la línea de investigación se vinculan temáticas que surgen de las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas, a fin de tratar con la optimización en problemas NP, o bien de poder trabajar en dominios de aplicación con soluciones integradas que hagan uso de técnicas y herramientas de estas disciplinas.

El trabajo de investigación se desarrolla en forma conjunta con investigadores afines de proyectos de esta Universidad, convenios entre organizaciones nacionales y provinciales con presencia en San Luis, como así también de universidades extranjeras mediante convenios de cooperación interinstitucional.

1. INTRODUCCIÓN

La optimización es una línea de investigación en Ciencias de la Computación, donde se procura encontrar la mejor solución posible a un problema dentro de un período de tiempo limitado. En el caso particular de problemas de optimización combinatoria, los hay NP-duros y los hay polinómicos; y no se puede garantizar encontrar la mejor solución en un tiempo razonable, para todas las instancias del problema.

En Geometría Computacional, la optimización de configuraciones geométricas respecto de ciertos criterios de calidad, pertenecen a esta clase de problemas, y pueden resolverse

utilizando métodos de aproximación, tales como las técnicas metaheurísticas [1].

Una metaheurística es un proceso de generación iterativo que guía la búsqueda de soluciones combinando inteligentemente diferentes conceptos de campos diversos como inteligencia artificial, evolución biológica, inteligencia colectiva, sistemas inmunes, entre otros [11].

Un objetivo de la línea es obtener soluciones aproximadas para problemas geométricos, para los cuales aún no se han encontrado algoritmos eficientes que los solucionen debido a su complejidad, mediante la aplicación de técnicas metaheurísticas. Las propuestas de estudio están orientadas a la optimización de diferentes configuraciones geométricas, tales como triangulaciones, pseudo triangulaciones, poligonizaciones, cuadrangulaciones, entre otras. Los criterios de calidad considerados son peso, dilación, número de apuñalamiento, número de guardias en problemas de vigilancia, área, perímetro, entre otros. Estos criterios inducen a la búsqueda de soluciones óptimas respecto de ellos, mediante la aplicación de técnicas metaheurísticas.

Algunos de los problemas de optimización estudiados son la Triangulación de Peso Mínimo (Minimum Weight Triangulation, MWT) y la Pseudo-Triangulación de Peso Mínimo (Minimum Weight Pseudo-Triangulation, MWPT), problemas de carácter NP-duro [9] [10]. La Triangulación de Dilación Mínima (Minimum Dilation Triangulation, MDT) es otro problema estudiado, donde la dilación mide la calidad de conexión entre puntos de la triangulación. Para este problema no se conoce un algoritmo que lo resuelva en tiempo polinomial y tampoco se ha demostrado que sea NP-duro.

Otro tópico de interés actualmente en análisis, refiere a problemas de vigilancia. Los problemas de vigilancia pueden interpretarse, muchas veces, como problemas de iluminación o vigilancia. Se han planteado numerosas variantes del problema, cuestionándose ¿qué se vigila? y ¿cómo se vigila? De esta forma, se trata con diversos objetos geométricos a vigilar y con diversas formas de vigilancia. En

particular, nos interesan como objeto geométrico de estudio las triangulaciones planas las cuales constituyen un entorno geométrico adecuado, en el que también tienen sentido las preguntas sobre vigilancia. Un guardia situado en un vértice (ó arista) de la triangulación vigila todos los triángulos incidentes al vértice (ó arista). Por tanto, dada una triangulación T , podemos considerar las siguientes preguntas ¿cuántos guardias, ubicados en vértices (ó aristas), se necesitan para vigilar todos los triángulos de T ?, y ¿dónde se deben ubicar? es decir, cuáles son los vértices (ó aristas) seleccionados para ubicar los guardias.

La utilización de estas configuraciones geométricas optimizadas respecto de algún criterio de calidad, resulta como soporte de estrategias en la resolución de problemas vinculados con bases de datos espaciales y espacio temporales. En este contexto, se propone el estudio y el desarrollo de herramientas para la visualización de aplicaciones vinculadas a las bases de datos mencionadas.

2. LÍNEA DE INVESTIGACIÓN

En la línea se investigan diversos dominios de aplicación de bases de datos espaciales y espacio temporales, con uso de técnicas y herramientas de apoyo en la resolución de problemas.

Entre los tópicos de estudio se pueden mencionar:

- Diseño y aplicación de índices espacio temporales, en diversos escenarios de movimiento (redes, espacios libres de obstáculos, etc.),
- Aplicación de la Geometría Computacional y su marco disciplinar, a los efectos de considerar aspectos propios de los problemas involucrados.
- Optimización de estructuras geométricas relacionadas con las bases de datos mencionadas, mediante la aplicación de metaheurísticas para la optimización de

problemas NP-duros en Geometría Computacional.

- Desarrollo de herramientas para la visualización de estructuras geométricas y aplicaciones vinculadas con bases de datos espacio temporales.
- Los procesos de descubrimiento para la recuperación y análisis de información oculta en grandes volúmenes de datos.

Como objetivos específicos en la línea de investigación se propone:

- Desarrollo de aplicaciones de Bases de Datos Espaciales y Espacio Temporales, con aplicación de herramientas de Geometría Computacional y la incorporación de procesos de descubrimiento de la información, utilizando métodos que permitan obtener información mediante análisis avanzados. Específicamente, el desarrollo de una plataforma prototipo, soporte para diversos eventos y sistemas de información con dominio de aplicación en el sector agropecuario, mediante el uso de las TIC. Las características principales están orientadas a la posibilidad de recolección de datos de diferentes fuentes, su almacenamiento compartido, la integración progresiva de diversas funcionalidades, la explotación y la visualización de la información. Dicha plataforma se propone disponible en la web, con accesibilidad mediante tecnología móvil (I+D+i).
- Estudio de la indexación espacio temporal sobre objetos en movimiento para diversos escenarios. Desarrollo de las estructuras de almacenamiento, algoritmos de consulta y evaluación experimental.
- Estudio de configuraciones geométricas de puntos en el plano considerando medidas de calidad, mediante la aplicación de técnicas metaheurísticas y diversas estrategias algorítmicas.
- Desarrollo de herramientas para la visualización de aplicaciones vinculadas a las bases de datos mencionadas.

3. RESULTADOS OBTENIDOS /ESPERADOS

En la resolución de problemas de optimización de triangulaciones y pseudo triangulaciones de peso mínimo se utilizaron diferentes técnicas metaheurísticas: Optimización basada en Colonia de Hormigas (Ant Colony Optimization, ACO) y Recocido Simulado (Simulated Annealing, SA), técnicas determinísticas Voraces (Greedy) y Triangulación de Delaunay. Se realizó la evaluación experimental de las técnicas metaheurísticas mencionadas para la búsqueda de triangulaciones y pseudo triangulaciones que cumplan la propiedad. Se diseñaron generadores de instancias de problema para ser utilizados en la evaluación experimental. Los resultados obtenidos para los problemas MWT y MWPT con la técnica ACO fueron publicados en [3] [7] [5]. Los resultados obtenidos con la técnica Recocido Simulado en [2] [4].

Para el problema Triangulación de Dilación Mínima las técnicas aplicadas fueron: Greedy, Local Search, Iterated Local Search, Simulated Annealing y Random Local Search. Para cada estrategia se propuso un conjunto de operadores adecuados. Se utilizó Optimización de Parámetros Secuencial (Sequential Parameter Optimization - SPO) para el ajuste de los parámetros requeridos por Simulated Annealing. Se realizó un análisis experimental en el cual se compararon dichos algoritmos con otras técnicas. Se crearon las instancias de prueba, ya que para estos problemas no se encontraron disponibles ningún tipo benchmark con el cual comparar los resultados. Se desarrolló un estudio estadístico aplicando diferentes test estadísticos y métodos de visualización [6].

Los resultados del tratamiento de estos problemas han dado lugar a tesis de doctorado.

También, se estudió el problema de obtener poligonizaciones de un conjunto de puntos con mínima área para un conjunto de puntos en el plano [13].

Por otra parte, se desarrolló una herramienta para la generación y visualización de

triangulaciones, pseudo triangulaciones y poligonizaciones de conjuntos de puntos en el plano [12]. Además, se implementó una aplicación en el ámbito de la Salud para el seguimiento de focos epidémicos utilizando base de datos espacios temporales y herramientas de Geometría Computacional [8]. El desarrollo de estas herramientas se plasmó en trabajos finales de Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas.

Por otra parte, con el fin de promover una plataforma de investigación, intercambio y desarrollo de tecnologías de la información y la comunicación (TIC), se integró el Proyecto Campo Conectado. Las organizaciones socias se proponen realizar acciones vinculadas a la promoción de las TIC en el ámbito de la producción agropecuaria con un horizonte inicial a dos años (2017-2018). El objetivo principal es aportar a la cooperación científico tecnológica y a las prácticas sociales, productivas y comerciales de los actores de la producción agropecuaria del semiárido central argentino.

En el marco de Campo Conectado, se propone accionar en la gestación de herramientas y desarrollos tecnológicos aplicados a la gestión de la producción agropecuaria en sistemas reales de producción. Actualmente, se está trabajando en el desarrollo de una herramienta para el seguimiento espacio temporal de rodeos en establecimientos agropecuarios, mediante el uso de las TIC y de las Bases de Datos Espacio Temporales, con herramientas de Geometría Computacional.

4. FORMACIÓN DE RECURSOS HUMANOS

La formación del grupo de trabajo en la Universidad Nacional de San Luis, se consolida con actividades de cooperación mutua e intercambio recíproco de información científica, tecnología y desarrollo de nuevos conocimientos con investigadores locales y de otras universidades.

Entre las actividades más destacadas, se mencionan:

- i) Formación de recursos humanos plasmada en tesis doctorales, tesis de maestría y Licenciados en Ciencias de la Computación.
- ii) Realización de pasantías de investigación con docentes de otras universidades.
- iii) Actividades de formación académica, a través del dictado y realización de cursos de posgrado y de especialización
- iv) Actividades de divulgación científica, conferencias y publicaciones en congresos y revistas en el ámbito nacional e internacional.
- v) Integración del proyecto interinstitucional Campo Conectado.

La línea tiene como un objetivo continuar con las actividades integradoras relacionadas al presente proyecto, proponiendo actividades de formación académica, de formación de recursos humanos locales y de otras universidades nacionales, investigación, desarrollo, y otras actividades académico-científicas vinculantes.

5. BIBLIOGRAFÍA

- [1] de Berg, M., Cheong O., van Kreveld, M., Overmars, M., *Computational Geometry: Algorithms and Applications*. 3rd edition, Springer-Verlag, Heidelberg, 2008.
- [2] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Approaches for MWT and MWPT Problems*. XVII Congreso Argentino de Ciencias de la Computación 2011 (CACIC 2011), 2011
- [3] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Using ACO metaheuristic for MWT problem*. XXX International Conference of the Chilean Computer Science Society. ISBN 978-0-7695-4689-6. Chile. 2011
- [4] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Metaheuristic approaches for MWT and MWPT Problems*. XIV Encuentros de Geometría Computacional. Pág: 79-82. 2011
- [5] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Approximations on Minimum Weight Triangulations and Minimum Weight Pseudo-Triangulations using Ant Colony Optimization Metaheuristic*. *Fundamenta Informaticae*. ISSN: 0169-2968 (Print), 1875-8681 (Online). Volume 119, number 1, Pag: 1-27.
- [6] Dorzán, M. G; Leguizamón, M. G.; Mezura-Montes, E.; Hernández Peñalver G. *Approximated algorithms for the Minimum Dilation Triangulation Problem*. *Journal of Heuristics*. DOI 10.1007/s10732-014-9237-2. Print ISSN 1381-1231. Online ISSN 1572-9397. Publisher Springer US. 2014.
- [7] Gagliardi E. O., Dorzán M. G., Leguizamón M. G. y Hernández Peñalver. G.; *Approximations on Minimum Weight Pseudo-Triangulation problem using Ant Colony Optimization*. XXX International Conference of the Chilean Computer Science Society. ISBN 978-0-7695-4689-6. Chile. 2011
- [8] Guasch, M.M; Piergallini, M.R; Dorzán, M.G.; Gagliardi, E.O.; Taranilla, M.T.; *“Una herramienta para el análisis y seguimiento de focos epidémicos”* en Anales del 17º Concurso de Trabajos Estudiantiles en 43 JAIO. Pág.: 35-45 Universidad de Palermo, Argentina, 2014.
- [9] Gudmundsson J., Levcopoulos C.; *Minimum weight pseudo-triangulations*. *Computational Geometry. Theory and applications*. Elsevier Vol. 38- Pages 139-153, 2007.
- [10] Mulzer W., Rote G. *Minimum weight triangulation is NP-hard*. In *Proceedings of the 22nd Annual ACM Symposium on Computational Geometry*. 2006.
- [11] Michalewicz Z., Fogel D., *How to Solve It: Modern Heuristics*, Springer, 2004.
- [12] Palmero, P.R., Dorzán, M. G., Gagliardi E.O., *Una Herramienta para la Manipulación de Configuraciones Geométricas*, 42º Jornadas Argentinas de Informática e Investigación Operativa - 16º Concurso de trabajos de fin de carrera. Universidad Nacional de Córdoba, Argentina. 2013. ISSN: 1850-2946.
- [13] Taranilla, M. T.; Gagliardi, E. O.; Hernández Peñalver, G. *Approaching Minimum Area Polygonization*. Congreso Argentino de Ciencias de la Computación (CACIC 2011), Pág.: 31-40, 2011.

APLICACIÓN DE MINERÍA DE DATOS PARA FACILITAR EL TRATAMIENTO DE LAS NORMAS DE PRODUCCIÓN DE ATRIBUTOS SEMÁNTICOS EN IDIOMA ESPAÑOL

Lucía Rosario Malbernat; Cecilia Ana Ruz; Silvia Adriana Cobialca
Grupo DM-AS, Departamento de Sistemas, Universidad CAECE

¹Sub sede Mar del Plata, ²Sede CABA

lmalbernat@ucaecmdp.edu.ar; cruz@caece.edu.ar; scobialca@caece.edu.ar

RESUMEN

Las Normas de Producción de Atributos consisten en un registro de aquellos aspectos compartidos de la memoria semántica referidos a la forma en que se definen los conceptos. La memoria semántica contiene información compartida por una comunidad de hablantes que les permite comunicarse.

La recolección de normas de producción de atributos semánticos ha proporcionado el insumo para la comprensión de numerosos fenómenos en el campo del estudio de la organización conceptual y la literatura científica señala la necesidad de contar con normas de atributos empíricamente derivados apropiadas para cada comunidad lingüística.

La recolección de las normas genera un volumen de datos muy rico en relaciones y correlaciones que, tratado con técnicas propias de la minería de datos pueden aportar reglas, patrones, tendencias y predicciones útiles para comprender las normas de producción de atributos y de interés para la Psicología Cognitiva y la Neuropsicología.

Se comparte en este trabajo la memoria técnica de un Proyecto de investigación recientemente aprobado en el marco de un Convenio de Cooperación en Investigación entre el Instituto de Psicología Básica, Aplicada y Tecnología. (IPSIBAT) de la Universidad Nacional de Mar del Plata que desde hace años define Normas de Producción de Atributos y la Universidad CAECE.

Palabras clave: *Data Mining*, *Big Data*, tendencias; patrones; predicciones; segmentaciones; clasificaciones; atributos

semánticos; normas de producción; lenguaje natural

CONTEXTO

El Proyecto que se reporta, “Aplicación de Minería de Datos para facilitar el tratamiento de las normas de producción de Atributos Semánticos en idioma español”, ha sido aprobado por R.R. 388/17 para el período 2018-2019 con fecha de inicio de actividades 1/04/2018 y tiene carácter inter sedes ya que participan investigadores de la Sede Central y la Subselección Mar del Plata.

Está radicado en el Departamento de Sistemas de la Universidad CAECE, donde se dictan entre otras, las Carreras Licenciatura e Ingeniería en Sistemas y Licenciatura en Gestión de Sistemas y Negocios y ha presentado recientemente para su aprobación una Maestría en Ciencias de datos e Innovación empresarial. Se va a desarrollar en el marco de actividades conjuntas en Investigación que se vienen realizando desde hace años, de acuerdo con convenios específicos, con el Centro de Investigación en Procesos Básicos, Metodología y Educación (CIMEPB) del Instituto de Psicología Básica, Aplicada y Tecnología (IPSIBAT) de la Universidad Nacional de Mar del Plata. Se propone analizar, procesar y modelar los datos generados por el grupo de investigación en Psicología Cognitiva y Educacional, especializado en la producción de normas de producción de atributos semánticos para el idioma español.

Entre las actividades conjuntas realizadas entre ambas instituciones se cuenta con la experiencia específica de haber aportado precedentemente *know how* y procesamiento

de datos, a través de técnicas específicas de minería de datos, mediante el proyecto, también radicado en el Departamento de Sistemas de la Universidad CAECE, “Aplicación de técnicas de *Data Mining* en gestión de docentes de educación superior” (DM-ES), aprobado por Resolución 549/13. El proyecto generador de los datos para dicho proyecto, radicado en el CIMEPB, era “Competencias para la innovación docente en enseñanza superior: preparación y actitud para el uso de las TIC” (Código 15/H2015; código de subsidio PSI221/14).

1. INTRODUCCIÓN

Las normas de producción de atributos semánticos consisten en colecciones empíricas de las características que las personas utilizan para describir cada concepto. Los datos se obtienen mediante una tarea de generación de propiedades que se logra pidiendo a los participantes, en un experimento controlado, que enumeren las características que mejor describen un cierto conjunto de conceptos. Esta tarea y las normas resultantes son relevantes en diversas áreas de la psicología, y se han utilizado durante décadas para resolver problemas teóricos y prácticos, pero solo recientemente algunas normas se hicieron públicamente disponibles [RVH04].

Establecer las normas de producción de atributos para un concepto permite obtener información cuantiosa acerca de las características y las relaciones de los conceptos y sus atributos para una población particular. Dada la importancia de los atributos semánticos para las teorías sobre memoria semántica, los investigadores han reconocido el valor de coleccionar normas de producción de atributos para construir modelos, testear hipótesis, disponer de estímulos experimentales y generar tareas de evaluación en el ámbito clínico [GRA04].

El volumen de datos generado en la recolección de normas de producción de atributos semánticos es cuantioso ya que implica interrogar a centenares de individuos sobre centenares de concepto. Dan cuenta de

dicho volumen, por ejemplo, la muestra tomada en el marco del Proyecto “Normas de Producción de Atributos Semánticos en Español Rioplatense en Adultos Mayores (Parte I) - 15/H247-” aprobado en 2017, conformada, tal como se describe en el artículo “*Spanish semantic feature production norms for 400 concrete concepts*” [VVC17] por 810 participantes de entre 20 y 40 años ($n=324.000$). De igual forma, la muestra con la que se trabajó durante la ejecución del Proyecto “Normas de producción de atributos semánticos en castellano rioplatense para un conjunto extenso de objetos vivos y no vivos.” (15/H178 , aprobado en 2011) referidos en el artículo “Distribución de los atributos semánticos en función del tipo de categoría y descripción del campo semántico” [VCG11] que reporta los resultados parciales de un sistema de normas de generación de atributos recogidos de 800 participantes sobre 400 conceptos referidos tanto a objetos vivos como a no vivos ($n=320.000$).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El Proyecto que se reporta en este trabajo tiene por objetivo **investigar la estructura** de los datos histórico recolectados por el grupo de investigación Psicología Cognitiva y Educacional para definir las normas de producción de atributos semánticos en idioma español, tomando la descripción de las características de las variables que manejan, dando cuenta, si corresponde, de la existencia de **errores o de datos atípicos** en la recolección de datos y de analizar la existencia de **posibles relaciones entre las variables** para finalmente aportar, -mediante la aplicación de técnicas funcionales y estructurales multivariantes y otras técnicas supervisadas y no supervisadas propias de la extracción de conocimiento y del procesamiento del lenguaje natural-, **predicciones, asociaciones, clasificaciones o segmentaciones de los datos** con el fin de resumir y visualizar la información de manera que se facilite la **identificación de tendencias o patrones** que los subyacen.

3. RESULTADOS ESPERADOS

Se ha propuesto recolectar todas las muestras que el Grupo de Investigación Psicología Cognitiva y Educacional viene tomando desde el inicio de sus proyectos vinculados con las normas de producción de atributos semánticos y en base a ellos **diseñar un modelo de datos** apropiado para el tratamiento de los datos mediante recursos de bases de datos y realizar a los datos recolectados todas las **transformaciones y recodificaciones** necesarias a modo de preparación inicial para su posterior análisis.

Se espera, así, poder **describir las características de las variables individuales** mediante distribuciones de frecuencia, medidas de tendencia central y medidas de variabilidad mediante estadística tradicional, analizar la posible **existencia de errores, datos ausentes o atípicos**, analizar la estructura interna de los datos mediante un análisis multivariado que permita detectar **tendencias y patrones**, aplicar técnicas de *Data Mining* para desarrollar reglas combinando las características, aplicando árboles de decisión, clasificaciones y segmentaciones y finalmente **armar un modelo predictivo** para predecir los conceptos.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación en Minería de Datos está integrado por profesores del Departamento de Sistemas que toman como insumo académico lo producido en el marco de los Proyectos de Investigación que desarrollan en el tema.

Anualmente se convoca a los estudiantes avanzados del Departamento para integrarse a los Proyectos de investigación aprobados. Recientemente, en el marco de las actividades del Grupo de Investigación en Minería de Datos, dos estudiantes han aprobado su Trabajo Final de Ingeniería (título del trabajo: “Evaluación de Validez de Esquemas de *Clustering* por Medio de la Aplicación de

Indices de Bondad”). Se espera que este año los estudiantes consideren como opción este nuevo Proyecto que el Grupo les pone a disposición.

5. BIBLIOGRAFIA

- [RVH04] Romero, C. Ventura, S & Hervás, C. Descubrimiento de Reglas de Predicción en Sistemas de e-learning utilizando programación genética. En: R. Giráldez, J. Riquelme & J. Aguilar-Ruiz. (Eds.) Tendencias de la Minería de Datos en España. Red Española de Minería de Datos 1 España, 2004
- [GRA04] Giráldez, R. Riquelme, J. & Aguilar-Ruiz, J. (Eds.) Tendencias de la Minería de Datos en España. Red Española de Minería de Datos. TIC2002-11124-E (ISBN 84-688-8442-1), 2004
- [RVP10] Romero C., Ventura, S. Pechenizkiy, M. & Baker R. *Handbook of Educational Data Mining*. Chapman and Hall/CRC Press, Taylor & Francis Group. ISBN: 9781439804575, 2010
- [KR02] Kimball, R. Ross, M. *The Data Warehouse Toolkit. The Complete Guide to Dimensional Modeling* (2a ed.) USA: Wiley Publishing, Inc, pp. 243-254, 2002
- [ARG09] Argibay, J.C. Muestra en investigación cuantitativa. Revista Subjetividad y Procesos Cognitivos, ISSN 1666-244X, N°. 13 p. 18, 2009
- [GRQ09] García Aretio, L. Ruiz Corbella, M. Quintanal Díaz, J. García Blanco, M. & García Perez, M. Concepción y Tendencias de la Educación a Distancia en América Latina. Colección Documentos de Trabajo. Centro de Altos Estudios Universitarios de la OEI, p. 44, 2009
- [VEG11] Vega Pons, S. Combinación de resultados de Clasificadores no supervisados. Tesis de doctorado. Rep. Téc. Reconocimiento de Patrones. Serie Azul. Cuba: Centro de Aplicaciones de Tecnologías de Avanzada, pp. 13, 47. 2011
- [YBP09] Yolis, E. Britos, P. Perichisky, G. & García-Martínez, R. Algoritmos Genéticos Aplicados a la Categorización Automática de Documentos. Revista Electrónica de sistemas de Información. ISSN 1677-3071 Doi: 10.5329/RESI, 2, 2009
- [BF98] Bradley P.S. & Fayyad, U.M. *Refining initial points for k-means clustering*. In J. Shavlik, editor, *Proceedings of the Fifteenth International Conference on Machine Learning (ICML '98)*, San Francisco, CA, 1998. Morgan Kaufmann, pp. 1-2, 91-99, 1998
- [NOR13] Northii, M. *Data Mining for the Masses. A Global Text Project Book* ISBN: 0615684378, 2013. Disponible en: <http://docs.rapid-i.com/files/DataMiningForTheMasses.pdf>
- [RAP12] *How to Extend RapidMiner 5 White Paper* Rapid-I GmbH, 2012
- [PSP12] *PSPP Users' Guide GNU PSPP Statistical Analysis Software Release 0.8.0-g13bf3f*, 2012. Disponible en: <http://www.gnu.org/software/pspp/manual/pspp.html>
- [BFH13] Bouckaert, R.R. Frank, E. Hall, M. Kirkby, R. Reutemann, P. Seewald, A. Scuse, D. *WEKA Manual for Version 3-7-10*, University of Waikato, Hamilton, New Zealand July 31, 2013
- [MVM17] MacIntyre, M. Vivas, L Vivas, J. *Tipologia de atributos ponderada baseadas em normas de produção atributos semânticos*. Temas em Psicologia 25 (2), 843-854, 2017
- [VVC17] Vivas, J. Vivas, L. Comesaña, A. Coni, A.G. Vorano. *Spanish semantic feature production norms for 400 concrete concepts*. *Behavior research methods* 49 (3), 1095-1106, 2017
- [VMR15] Vivas, J. MacIntyre, M. Ricci, L. Vivas, J. *Psycholinguistic variables involved in concept recall from the successive presentation of features*. *Estudios de Psicología* 36 (3), 592-619, 2015
- [CV15] Comesaña, A. Vivas, J. Evolución de la categorización semántica en adultos mayores con diagnóstico de DCL-A y DTA y sin patología neurológica Interdisciplinaria 32 (1), 7-29, 2015
- [CSA14] Cervigni, M Sguerzo, M.R. Alfonso, G. Pastore, M. Martino, P. Mazzoni, C. *Bibliometric analysis of empirical studies in Spanish on Working Memory* (1999-2014), 2014
- [ZV15] Zapico, M. Vivas, J. La sinonimia desde una perspectiva lingüístico-cognitiva. Medición de la distancia semantic. *Onomázein*, 198-211, 2015
- [HLV15] Huapaya, C. R. Lizarralde, F.A.J. Vivas, J. Modelo para visualizar y evaluar el conocimiento conceptual. TE & ET 2015

- [FCM14] Favarotto, V Coni, A.G. Magani, F. Vivas, J. *Semantic Memory Organization. In Children And Young Adults. Procedia-Social and Behavioral Sciences* 140, 92-97, 2014
- [ZV14] Zapico, M. Vivas, J. *La sinonimia como caso particular de distancia semantic. Encontros Bibli: revista eletrônica de biblioteconomia e ciência da*, 2014
- [LVV12] Lamas, V Vivas, J. Vorano, A. Comparación de atributos semánticos entre diferentes lenguas. IV Congreso Internacional de Investigación y Práctica Profesional en Psicología, 2012
- [MVF12] Morales, F. Vivas, J. Vorano, A. Intyre, M. Damian, M. Campoy, P. Normas de Producción de Atributos Semánticos: Diferencias de acuerdo a la edad. IV Congreso Internacional de Investigación y Práctica Profesional en Psicología, 2012
- [VCG11] Vivas, J. Comesaña, A García Coni, A Vivas, L. Yerro, M. Distribución de los atributos semánticos en función del tipo de categoría y campo semántico. Resultados preliminares para la confección de normas de atributos. M.C. Richaud y V. Lemos (comp.) *Psicología y otras ciencias del comportamiento*, 2011
- [PYF11] Pazgón, E. Yerro Avincetto, M Favarotto, V. Vivas, L. Vivas, J. Categorización de rasgos semánticos: Diferencias de género en una tarea de atributos de conceptos. *Perspectivas en Psicología: Revista de Psicología y Ciencias Afines* 8 (2), 2011

Tecnologías de la Web Semántica aplicadas al tratamiento de documentos jurídicos electrónicos

Héctor J. Ruidías^{1a}, Karina B. Eckert^{1b}, Juan M. Lezcano^{2c}, Carolina V. Rosas^{2d}

¹Departamento de Ingeniería y Ciencias de la Producción

²Departamento de Ciencias Sociales y Jurídicas

Universidad Gastón Dachary - Posadas, Misiones, Argentina

{^achandra149, ^bkarinaeck, ^cjuanmanuellezcano, ^dcarolinav.rosas}@gmail.com

RESUMEN

En el terreno jurídico, el crecimiento de Internet ha permitido que las normativas legales, digestos, fallos judiciales, sentencias y leyes estén a disposición a través de portales de acceso público o privado, los cuales cuentan con una forma de organización determinada y búsqueda por términos. La documentación jurídica supone aspectos intrínsecos del objeto judicial que posee relaciones de categorías propias del derecho y de las legislaciones existentes que muchas veces no aparecen en forma explícita y que requieren por lo tanto sean identificadas en un proceso de “curación” del documento legal.

Se presenta así la necesidad de ofrecer una forma de tratamiento documental que ponga atención a la semántica de los textos. Una forma de abordar esta cuestión es a través de las tecnologías semánticas, en particular recurriendo al uso de *ontologías* que ofrecen una forma *explícita* de representación que permite aplicar procesos de *descubrimiento* de relaciones a partir de mecanismos de razonamiento automático. En este modelo de relaciones semánticas explícitas, los mecanismos de búsquedas pueden resultar favorecidos al permitir búsquedas sobre un vocabulario menos dependiente

de la simple ocurrencia terminológica, y sí más orientado a las estructuras de las categorías empleadas (o de los metadatos y sus relaciones).

Palabras clave: Gestión de documentos legales, Web Semántica, Ontologías, Razonamiento automático, Gestión del Conocimiento.

CONTEXTO

La presente investigación se encuentra en ejecución, acreditado en la Secretaría de Investigación y Desarrollo de la Universidad Gastón Dachary (UGD) por resolución N° 07/A/17. El proyecto es del tipo interdisciplinario abarcando dos áreas disciplinarias principales: Informática (Tecnologías de la Información y Comunicación (TIC)) y Ciencias Jurídicas (Derecho).

1. INTRODUCCIÓN

El fenómeno de la inflación legislativa y la producción de documentos jurídicos hacen necesario desarrollar y promover plataformas y medios de comunicación e información que proporcionen oportunidades de acceder, compartir e intercambiar los recursos científicos, culturales, sociales y económicos que

están disponibles en base de datos vinculadas con las ciencias jurídicas [1].

Atento a que dicho fenómeno se da también en un contexto de crecimiento exponencial de la información producida globalmente por organizaciones, empresas e individuos, alentada sin duda por la dinámica propiciada por Internet, es precisamente que también sobre la misma coyuntura científico-tecnológica se están desarrollando el mayor número de avances tecnológicos que abordan dicha problemática desde dos aspectos fundamentales: uno relacionado al tratamiento de volúmenes masivos de datos y el otro relacionado a la reducción de la opacidad semántica de dichos datos.

Es por ello que en forma concomitante al desarrollo de la Web, tal como fuera planteada por Sir Tim Berners Lee, también ha ido tomando vigor la idea de la Web Semántica [2], donde la información no solamente resultara comprensible por seres humanos, sino también por máquinas.

El hecho de que la búsqueda de información jurídica sea deficiente, no resulta únicamente en un inconveniente de índole técnico, sino que atenta contra los mismos cimientos de la democracia a propiciar la falta de transparencia y celeridad necesaria en la resolución de casos judiciales que esconden un entramado de conflictos sociales [3], [4].

Las ontologías han sido propuestas como artefactos de representación que especifican un vocabulario relativo a cierto dominio en el contexto de los sistemas de información [5]. Dicho vocabulario define entidades, clases, propiedades, axiomas y reglas, además de las relaciones entre dichos componentes, con el objetivo de reducir la ambigüedad, los conflictos terminológicos y las discrepancias semánticas que se presentan en un área de conocimiento determinada. Esto conduce a una definición donde una

ontología es entendida como “una especificación formal y explícita de una conceptualización compartida” [6]. Se destaca en la misma dos ideas claves, la de conceptualización compartida y la de especificación formal. La primera refiere a un modelo abstracto, con sus elementos y la vinculación entre estos, explícitamente definido y además cumple con el requisito de ser consensuada y aceptada por una comunidad. El segundo en tanto aborda la cuestión del lenguaje y el vocabulario con que se representa, de tal manera que sea interpretable de igual manera tanto por personas como por computadoras, siendo por lo tanto una cualidad irrenunciable el que sea procesable por máquina.

El énfasis en el aspecto “consensuado” que propician las tecnologías de la Web Semántica va desde la estandarización y especificaciones, la arquitectura de capas de la W3C, hasta las definiciones de ontologías de alto nivel (o fundacionales) o específicas para distintos dominios, ya sean médicos, de organizaciones y personas o de índole legal. A modo de resumen se puede mencionar el conjunto de tecnologías de base sobre las que se construye la Web actual, que incluye una forma unívoca de identificar recursos (URI), una codificación estándar (UNICODE), y un lenguaje básico para estructurar documentos (XML+XMLSchema) y sobre el cuál se respaldan los distintos lenguajes de marcado, tales como XHTML y HTML. Sobre tales tecnologías se cimientan las tecnologías específicas de la web semántica, que incluyen descriptores de recursos y relaciones entre estos (RDF y RDFSchemas) y un lenguaje específico de ontologías (OWL), que aporta mayor expresividad al permitir describir propiedades y clases con restricciones de cardinalidad, exclusión (relaciones disjoint), y una mayor riqueza a la hora de

definir propiedades (relaciones). Naturalmente tales características imponen algunas restricciones en función de la expresividad deseada, lo cual permite diferenciar entre tres tipos básicos de familias de ontologías, OWL Lite, OWL DL y OWL FULL [7].

Las ontologías han demostrado su utilidad desde sus orígenes al ofrecer un marco conceptual útil para reducir la ambigüedad inherente al uso del lenguaje natural. Puesto que en muchos dominios del conocimiento humano se establecen definiciones en términos de construcciones verbales susceptibles a varias interpretaciones, principalmente aquellas que se vinculan a las ciencias sociales.

En el derecho, la interpretación es un eje fundamental sobre el que se articula el accionar aplicativo de la justicia, y el deterioro legislativo debido a la sobresaturación de leyes y normas, puede suponer un reto a la hora de conocer la vigencia real sobre cierta normativa. Todo esto pone sobre relieve la necesidad de establecer un proceso de ordenamiento jurídico [3] que culmina en la creación de un Digesto Jurídico. Tal proceso supone básicamente el escrutinio permanente sobre el crecimiento legislativo y el consecuente ordenamiento del mismo. Existen tres categorías fundamentales de intervención que se consideran y que definen la complejidad del proceso (consolidación), que incluye 1) un aspecto sistemático, es decir la forma en que se organizará la normativa, 2) otro lingüístico (morfológico, sintáctico y semántico) y 3) uno de interrelación normativa.

Estas tres categorías resultan de vital importancia, no únicamente para el ordenamiento jurídico, sino para la digitalización y su tratamiento a través de sistemas de gestión documental, que

permitan almacenar y acceder a dicha normativa.

La creación de un Digesto Jurídico por otra parte supone la conformación a una doctrina, un marco jurídico y normativo y también el entrecruzamiento entre las ramas clásicas del derecho y otras transversales tales como “equidad de género”, “biodiversidad”, “derecho a la salud” entre otras [4].

En el marco de un sistema de gestión documental enriquecido con tecnologías semánticas, un Digesto Jurídico se define a partir de vinculaciones semánticas (anotado semántico) entre el texto de la norma y una ontología de referencia. Para ello lo apropiado supone la reutilización de ontologías de alto nivel [8], [9] para objetos que no son propios del derecho, tales como aquellos que refieren a eventos temporales y ubicaciones espaciales, y para aquellos que pertenecen a categorías propias del Derecho, emplear ontologías de dominio tales como FOLAW o LRI-CORE [10], [11].

Un antecedente destacado que se puede mencionar es el e-COURT de la Unión Europea, un sistema para manejo de información legal, textual y multimedia, que permite el almacenamiento y recuperación de la misma, y que es tolerante a ambigüedades de índole lingüística gracias a las tecnologías semánticas que emplea [12]. Otro antecedente es el Digesto Jurídico basado en ontologías creado para la Universidad Nacional del Litoral [13].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La presente línea de investigación comprende aspectos propios del derecho, como el tratamiento y ordenamiento jurídico de documentos legales, y otros propios de las TIC, como ser todo lo

relativo a la Web Semántica y tecnologías relacionadas, así también como sus aspectos metodológicos (Ej. Metodologías para desarrollar ontologías: methontology). Por otra parte se emplea documentos legales reales para tener una mejor dimensión de la problemática, en base a los cuales se diseña y pone a prueba una infraestructura de software desarrollada a partir de distintas tecnologías de la web semántica y sistemas de base de datos no relacionales y/o documentales. Cabe mencionarse que la propuesta no se centra únicamente en el hipotético sistema de una respuesta respecto a qué documentos recuperar a partir de un término de búsqueda, sino también a un dato puntual solicitado (Ej. todos los fallos vinculados a “violencia de género” que fueron “apelados”), eventualmente en lenguaje natural, aunque podría formularse a través de una interfaz ad-hoc similar al lenguaje de consulta SPARQL.

El trabajo desarrollado sentará las bases para la transferencia de soluciones tecnológicas al sector judicial, partiendo desde un prototipo funcional que empleará datos aportados desde el Poder Judicial de la Provincia de Misiones mediante la firma de un convenio de transferencia de conocimientos entre la UGD y dicho organismo estatal. Asimismo desde el punto de vista académico, desde el trabajo llevado a cabo con tecnologías semánticas, se puede ver ya no solo un campo fértil para la aplicaciones de tales tecnologías, sino también la posibilidad de acceder a una fuente de escenarios de prueba para la validación de las propuestas desarrolladas tanto en el marco del presente proyecto como de futuros proyectos.

3. RESULTADOS Y OBJETIVOS

Objetivo General

Proponer una infraestructura orientada a la búsqueda de información jurídica en documentos electrónicos legales a partir de la integración de tecnologías semánticas.

Objetivos Específicos

- Establecer un marco teórico-conceptual sobre los antecedentes de la problemática en torno al ordenamiento jurídico de leyes, fallos, sentencias y otros tipos de documentos jurídicos, como así también de los medios manuales o mecanizados para acceder a los mismos, principalmente enfocado en los sistemas informatizados y de búsqueda por ocurrencia de términos ya sea de coincidencia plena o aproximada (ya sea por criterios estadísticos o semánticos).
- Comprender la relación entre términos de búsqueda, tesauros, digestos jurídicos y ontologías aplicadas al ámbito jurídico.
- Determinar la pertinencia de distintas soluciones tecnológicas de base de datos no relacionales para el almacenamiento y la recuperación de documentos judiciales con tecnologías semánticas.
- Especificar una ontología integrada con ontologías de alto nivel y específicas del dominio para la gestión de documentos jurídicos.
- Proponer una solución unificada de organización de la información jurídica en relación a los las estrategias de búsqueda semántica.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está conformado de la siguiente manera:

Director del Proyecto:
 Ing. Héctor Javier Ruidías
 Co-Director:
 Dr. Juan Manuel Lezcano
 Docentes-Investigadores:
 Ing. Karina Beatriz Eckert
 Abogada Carolina Vanesa Rosas
 Tutor/Asesor
 Dra. María L. Caliusco

El Ing. Héctor J. Ruidías se encuentra desarrollando su tesis doctoral sobre tecnologías semánticas en UTN Facultad Regional Santa Fe bajo la dirección de la Dra María L. Caliusco.

Además el equipo incluye a estudiantes de grado de las carreras de Ingeniería en Informática y Derecho quienes colaboran en diversas tareas.

5. REFERENCIAS

- [1] G. L. Paul y J. R. Baron, “Information Inflation: Can the Legal System Adapt?”, *Rich JL Tech*, vol. 13, pp. 10–17, 2007.
- [2] T. Berners-Lee, J. Hendler, O. Lassila, y others, “The semantic web”, *Sci. Am.*, vol. 284, no 5, pp. 28–37, 2001.
- [3] R. G. Brenna, “El Digesto Jurídico. Un camino hacia la democratización de la información jurídica”. [En línea]. Disponible en: <http://www.dab.com.ar/articles/53/eldigesto-jur%C3%ADdico-un-camino-hacia-la-democratiza.aspx>. [Accedido: 14-nov-2017].
- [4] M. A. Ciuro Caldani, “EL COMPLEJO DEL FUNCIONAMIENTO DE LAS NORMAS”, *INVESTIGACIÓN Y DOCENCIA*, vol. 40, 2007.
- [5] N. Guarino, *Formal Ontology in Information Systems: Proceedings of the First International Conference (FOIS'98)*, June 6-8, Trento, Italy IOS Press, 1998.
- [6] T. R. Gruber, “A translation approach to portable ontology specifications”, *Knowl Acquis*, vol. 5, pp. 199–220, jun. 1993.
- [7] T. Berners-Lee, “Semantic Web on XML”. [En línea]. Disponible en: <http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide11-0.html> [Accedido: 26-2-2018].
- [7] K. Kaneiwa, M. Iwazume, y K. Fukuda, “An Upper Ontology for Event Classifications and Relations”, en *AI 2007: Advances in Artificial Intelligence*, M. A. Orgun y J. Thornton, Eds. Springer Berlin Heidelberg, 2007, pp. 394–403.
- [8] V. Mascardi, A. Locoro, y P. Rosso, “Automatic Ontology Matching via Upper Ontologies: A Systematic Evaluation”, *IEEE Trans. Knowl. Data Eng.*, vol. 22, no 5, pp. 609–623, may 2010.
- [9] J. a. P. J. Breukers y R. J. Hoekstra, “Epistemology and ontology in core ontologies: FOLaw and LRI-Core, two core ontologies for law”, *CEUR Workshop Proc.*, 2004.
- [10] R. Benjamins, *Law and the Semantic Web - Legal Ontologies, Methodologies, Legal Information Retrieval, and Applications*. Springer, 2005.
- [11] J. Breuker, A. Elhag, E. Petkov, y R. Winkels, “Ontologies for legal information serving and knowledge management”, en *Legal Knowledge and Information Systems, Jurix 2002: The Fifteenth Annual Conference*, 2002, pp. 1–10.
- [12] F. Enrique y M. Lopez, “Desarrollo de algoritmos de procesamiento para la indexación y la búsqueda en lenguaje natural del contenido de un digesto basado en tecnologías semánticas”, *Universidad Nacional del Litoral*, 2016.

EXTRACCIÓN DE CONOCIMIENTO EN REDES SOCIALES MEDIANTE HERRAMIENTAS DE SOFTWARE LIBRE Y PLATAFORMAS DE HARDWARE PARALELO-DISTRIBUIDAS

Gouiric, Guillermo Adrián; Ortega, Manuel Oscar; Klenzi, Raúl Oscar
Instituto de Informática / Departamento Informática / Facultad de Ciencias Exactas
Físicas y Naturales / Universidad Nacional de San Juan
Domicilio: Av. Ignacio de la Roza 590 (O), Complejo Universitario "Islas Malvinas",
Rivadavia, San Juan, CPA: J5402DCS, 0264-260353 0264-4260355
{guillegouiric;manuel.ortega;rauloscarklenzi}@gmail.com

RESUMEN

El presente trabajo pretende realizar tareas de extracción de conocimiento en grandes colecciones de datos mediante la aplicación de la herramienta de software libre de aprendizaje de máquina KNIME ANALYTICS procesando datos provenientes de la Red Social (Twitter) y ejecutándose en plataformas paralelo distribuidas tratando de cotejar las mejoras de performance respecto de las aplicaciones secuenciales en la caracterización de perfiles de usuario. A tal efecto se habrá de trabajar con la versión de KNIME ANALYTICS 3.5.2, ejecutándose sobre un cluster de cuatro terminales de cómputo constituyendo el paradigma de computación distribuida cada una de las cuales cuenta con placas GPU computing sobre una de las cuales se ejecutará la instancia paralela del análisis.

Palabras clave: *Minería de Datos, Redes Sociales, Paralelismo, HPC, GPU.*

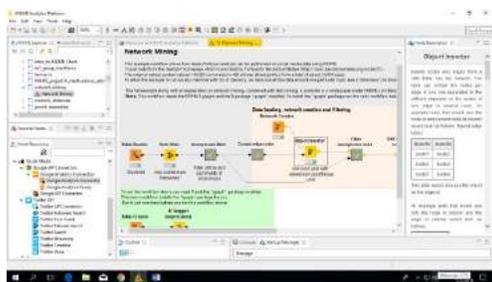
CONTEXTO

En el ámbito del proyecto CICIPCA_UNSJ "Ciencia de los Datos

aplicada a grandes colecciones de datos" ejecutado en el bienio 2016-2017 y la continuidad buscada por el grupo de investigadores que lo conformaban, en la presentación de un proyecto para el bienio 2018-2019 actualmente en evaluación "Visualización y DeepLearning en Ciencia de los Datos" se ha conformado un cluster de cuatro?? Computadoras cada una de las cuales cuenta con unidades NVIDIA de GPU computing. Con el hardware citado y desde la realización de trabajos finales de grado y becas de investigación de alumnos avanzados se lleva adelante la determinación de perfiles de usuarios de TWITTER y Redes sociales en general, mediante la utilización de software libre de aprendizaje de máquina y cómputo paralelo distribuido.

La elección del entorno de software KNIME ANALYTICS 3.5.2 se centra en lo expresado en [1] KNIME es una plataforma cohesionada para científicos de datos de todos los niveles de habilidades, que proporciona un marco de ciencia de datos único y consistente. Ofrece capacidades de acceso y manipulación de datos de alta calificación, una amplia y completa gama de algoritmos y herramientas de aprendizaje automático adecuadas tanto

para principiantes como para científicos de datos experimentados. La plataforma de KNIME se integra con otras herramientas y plataformas, como R, Python, Spark, H2O.ai, Weka, DL4J y Keras, a la vez que permite su ejecución considerando procesadores multinúcleos y/o GPGPU computing. La ayuda contextual de KNIME es más flexible que los "asistentes" fijos. La interfaz de usuario y los extensos ejemplos proporcionados con la plataforma atraen a la comunidad de científicos de datos.



Interfaz de KNIME ANALYTICS 3.5.2

1. INTRODUCCIÓN

Para explicitar conceptos brindados en el contexto de la propuesta, habrán de definirse aspectos relevantes a la misma:

Big Data hace referencia a una colección de datos de considerable dimensión que hacen imposible el procesamiento con aplicaciones tradicionales de base de datos.

En [2] definen *Big Data* usando las tres V's: Volumen, Velocidad y Variedad.

Volumen se refiere a la cantidad de datos, desde Terabytes (TB) a Petabytes (PB), relacionado con la estructura de estos datos incluyendo registros, transacciones, archivos, y tablas. El

volumen de estos datos, se prevé que crecerá 50 veces para el 2020.

Velocidad se refiere a la forma de transferir los grandes volúmenes de datos incluyendo transmisión en batch, tiempo real y flujos. La velocidad, incluye tiempo y latencia, características propias del manejo de datos. Los datos pueden ser analizados, procesados, almacenados y manipulados en forma rápida, o con un retardo entre eventos.

Variedad de los grandes volúmenes de datos, se refiere a los diferentes formatos que pueden adoptar los datos, incluyendo estructurados, semi estructurados, desestructurados y todas las combinaciones de estos tres. El formato de los datos incluye: documentos, mails, mensajes de texto, audio, imágenes, video, gráficos, entre otros.

High-Performance Computing (HPC) según [3] se usa para describir ambientes de cómputo que utilizan supercomputadoras o clusters de computadoras para atender requerimientos de cómputo complejos o aplicaciones con requerimientos altos de tiempo o que requieren procesamiento de grandes volúmenes de datos.

La tecnología HPC es apropiada tanto para las aplicaciones de cálculo intensivo, como las de procesamiento intensivo de datos. Las plataformas HPC utilizan un alto grado de paralelismo que tiende a usar multiprocesadores especializados con arquitecturas de memoria que han sido altamente optimizadas para cálculos numéricos.

Las PC's actuales tienen más poder de cómputo que las supercomputadoras de hace una década. Estas PC's, poseen procesadores con múltiples cores (multicore), con poderosas placas de video (GPGPU), inicialmente concebidas para abordar tareas gráficas que requieren gran cantidad de cómputo en paralelo. Las PC's con procesadores multicore, y placas gráficas GPU, constituyen lo que en [4] define como computación heterogénea, ya que combina más de un tipo de procesador. Este tipo de configuración de hardware, constituye una herramienta nueva para la computación HPC.

Redes sociales

En las redes sociales son sitios de Internet formados por comunidades de individuos con intereses o actividades en común (como amistad, parentesco, trabajo) y que permiten el contacto entre estos, de manera que se puedan comunicar e intercambiar información.

Minería de Datos

La minería de datos, Data Mining, según [6] es un proceso de descubrimiento de nuevas y significativas relaciones, patrones y tendencias al examinar grandes cantidades de datos. La disponibilidad de grandes volúmenes de información y el uso generalizado de herramientas informáticas ha transformado el análisis de datos orientándolo hacia determinadas técnicas especializadas englobadas bajo el nombre de minería de datos o Data Mining. Las técnicas de minería de datos persiguen el descubrimiento automático del conocimiento contenido en la información almacenada de modo ordenado en grandes bases de datos.

Sistemas Distribuidos

Los sistemas distribuidos intentan hacer que un conjunto de computadoras físicamente separadas (cada una con sus propios recursos como procesador, memoria, buses) trabajen cooperativamente para resolver un problema grande, y que externamente se vean como una sola unidad con un gran poder de procesamiento.

Se define en [7] un Sistema Distribuido como un sistema en el cual componentes de hardware y software, localizadas en computadores de red, se comunican y coordinan sus acciones sólo por paso de mensajes

Las computadoras conectadas mediante red pueden estar físicamente a cualquier distancia, inclusive separadas por continentes, pero también en el mismo edificio o habitación.

CUDA-Paralelismo

Según [8] a través del uso de CUDA es posible construir aplicaciones paralelas capaces de aprovechar los múltiples cores de una GPU. Hacer programas para ser ejecutados sobre una GPU implica comprender la interacción de datos y de control que se produce entre la CPU y la GPU, dando lugar a una nueva modelo de programación paralela. Este modelo combina un enfoque de trabajo distribuido con uno de trabajo paralelo. El enfoque distribuido es consecuencia del hecho de que la CPU y la GPU tienen memorias disjuntas y arquitecturas de hardware diferentes. Además, ejecutan tareas que, si bien están relacionadas, no son las mismas, y por lo tanto no ejecutan las mismas instrucciones. El enfoque paralelo se da por el uso de los múltiples cores de la GPU, de modo que todos pueden ejecutar la misma instrucción al mismo tiempo, pero sobre diferentes datos

Rendimiento o performance

El rendimiento es la rapidez con la que el computador puede ejecutar programas, es la inversa del tiempo requerido por una computación, el cual se calcula con la siguiente expresión:

$$\frac{1}{\text{Tiempo de ejecución}}$$

Speed Up

Es una métrica usada para medir la mejora relativa de la performance de un programa cuando a este se le hace alguna modificación. La forma básica para calcular el Speed Up es:

$$\frac{\text{Tiempo de ejecución sin la mejora}}{\text{Tiempo de ejecución con la mejora}}$$

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el marco de la investigación se pretende:

- Analizar y describir el conjunto de estudios y prácticas requeridos en Ciencia de Datos.
- Analizar diferentes API's (Application Programming Interface) de aplicación para el relevamiento de datos Redes Sociales y su análisis e interpretación, sobre plataformas paralelo-distribuidas, con las formas de representación que las herramientas a utilizar poseen.

3. RESULTADOS ESPERADOS

Se espera comprobar mejoras en la performance, de los algoritmos paralelo distribuidos respecto de las alternativas secuenciales. Especialmente se espera una interesante mejora en el SpeedUp de las aplicaciones distribuidas y fundamentalmente paralela. Se trabajará centralizando las actividades en la interfaz de KNIME dado que la misma tiene capacidades de interactuar con

diferentes entornos de software y trabajar en plataformas multinúcleos y/ GPGPU computing.

4. FORMACIÓN DE RECURSOS HUMANOS

Esta propuesta, originalmente contenida en el proyecto “Ciencia de los Datos en Grandes colecciones de Datos”, permite continuarse en el proyecto actualmente en evaluación “Visualización y DeepLearning en Ciencia de Datos”. En este contexto la profundización del conocimiento adquirido por los integrantes del proyecto, permite formar nóveles docentes-investigadores del departamento informática, así como la dirección y defensa de diferentes tesis de maestría y trabajos finales de grado, Simultáneamente, permite proponer alumnos de grado a becas de alumnos avanzados e iniciación. La significativa sinergia alcanzada por el grupo de trabajo ha permitido extender las aplicaciones del Data Science a otros proyectos a la vez que proponer diferentes instancias de propagación del conocimiento que se vuelca a las cátedras de las carreras del Departamento Informática, así como en charlas, y/o cursos de posgrado.

5. BIBLIOGRAFÍA

- [1] «Gartner 2018 Magic Quadrant for Data Science and Machine Learning – Read the report». [En línea]. Disponible en: <https://www.kdnuggets.com/2018/02/domino-gartner-mq-data-science-machine-learning.html>. [Accedido: 16-mar-2018].
- [2] B. Furht y F. Villanustre, *Big Data Technologies and Applications*. Springer, 2016.
- [3] B. Furht y A. Escalante, *Handbook of Data Intensive Computing*.

- Springer Science & Business Media, 2011.
- [4] Y. Tan, *GPU-based Parallel Implementation of Swarm Intelligence Algorithms*. Morgan Kaufmann, 2016.
- [5] Anónimo.(s.F). *Concepto de Redes Sociales*. Recuperado de <http://concepto.de/redes-sociales/>
- [6] GestioPolis.com Experto. (2001); *¿Qué es Data Mining?* Recuperado de www.gestiopolis.com/que-es-data-mining/
- [7] G. Coulouris, *Sistemas distribuidos, conceptos y diseño*, Addison Wesley, 2002.
- [8] N. Wilt, *The CUDA Handbook*, Addison-Wesley, 2013.
- .

IMPLEMENTACIÓN DE TÉCNICAS ESPECÍFICAS DE MINERÍA DE DATOS EN APLICACIONES WEB CON MOTORES DE BASE DE DATOS RELACIONALES

VARGAS, Luis Alejandro; FARFAN, José Humberto; RODRIGUEZ, Mariela Ester; TAPIA, Marcela Alejandra; PAREDES, Julio Cesar RAMOS, Pablo Nicolás; LLAMPA, Alvaro Facundo; MONTES, Leonardo Ezequiel; MOGRO, Nelson Ariel & CORDOBA, Irma Rafaela Mercedes;

Ingeniería + Software (I+S), Área de Ingeniería Informática, Facultad de Ingeniería, Universidad Nacional de Jujuy (U.N.Ju).

ARAMAYO, Fernando Rubén & SPADONI, Gustavo Fernando
Facultad de Ciencias Agrarias, Universidad Nacional de Jujuy (U.N.Ju.)

RESUMEN

El presente proyecto pretende implementar técnicas de Datamining o también denominado Minería de Datos en aplicaciones web de sistemas informáticos y procesan la información con motores de Base de Datos Relacionales, es decir efectuar “Análisis Inteligente de Datos”. En el desarrollo de tecnologías web genera la necesidad de contar con herramientas, carencia que se ha analizado y detectado en investigaciones anteriores, por lo que se pretende implementarlo en lenguajes específicos de programación web. Implica obtener ventajas y beneficios de la Minería de Datos ya sea aplicando técnicas específicas de descripción o de aprendizaje automático. Se puede mencionar Redes Neuronales, reglas de asociación, clustering o similares, en cualquier problema genérico de un Sistema Informático desarrollado en un lenguaje de programación web. El objetivo es realizar en los datos un “Análisis Inteligente de Datos”. Para lograr este objetivo es necesario clasificar y estudiar en profundidad las técnicas que representan a la Minería de Datos,

aplicándolas en lenguajes de programación web e implementarlos con motores de Base de Datos. Es obligatorio y necesario conocer en detalle los lenguajes involucrados, estudiar la implementación correcta de las técnicas de Minería de Datos y la conectividad con el motor de Base de Datos.

Palabras Clave: Datamining, Análisis Inteligente, Programación Web

CONTEXTO

El proyecto se encuentra inserto dentro de las siguientes Líneas Prioritarias de Investigación de la Facultad de Ingeniería (LIPIFI) - UNJu:

- Ingeniería del Software
- Ingeniería de Procesos

Es un proyecto aprobado de categoría B (Código D/B030)

Financiamiento: Secretaría de Ciencia y Técnica y Estudios Regionales (SeCTER) de la UNJu.

Vigencia del Proyecto: 01/01/2018 al 31/12/2019

1. INTRODUCCIÓN

Se entiende que la Minería de Datos es “un campo de la estadística y las ciencias de la computación referido al proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos” [1], el actual Grupo de Investigación se inició en el año 2.016 con el proyecto, aprobado, de categoría B denominado “Data Mining aplicado a análisis telefónico”, el mismo ha finalizado en diciembre de 2017, siendo importante destacar que se aplicaron técnicas de Minería de Datos exclusivas para encontrar soluciones al problema en cuestión. Este último es un problema genérico en el que se realizó un Análisis de Datos de una Red de Comunicaciones con el formato actual que proveen las compañías telefónicas denominado “Sábanas de llamadas” para ser aplicados en el ámbito de la Seguridad de Organismos Gubernamentales Provinciales y/o Nacionales, siendo necesario por ello examinar y describir técnicas y herramientas que emergen en esa área de investigación aplicadas a la toma de decisiones [2]. El descubrimiento de la información oculta es de importancia estratégica, y es posible por las características de la Minería de Datos, pero es el descubrimiento del conocimiento (KDD, por sus siglas en inglés) el que se encarga de la preparación de los datos y de la interpretación de los resultados obtenidos, los cuales darán significado a los patrones encontrados [3]. Se destaca que KDD es producto del rápido desarrollo de la minería de datos y la aplicación de tecnologías de información y bases de datos. Zhang et al [4] formulan el proceso basado de

la extracción de conocimiento (KDD) en una secuencia iterativa de cuatro pasos: la definición del problema, el pre - procesamiento de datos (que incluye la preparación de datos), data mining, y el post data mining.

Como se mencionó anteriormente el proyecto de investigación “Data Mining aplicado a análisis telefónico” se encuentra en su última fase, centrado en el análisis y aplicación de técnicas específicas de minería de datos al problema de estudio, abarcando y estudiando los resultados que son de utilidad a los usuarios finales que hacen uso de la misma, principalmente en el área de seguridad gubernamental.

Sin embargo existe una cantidad de técnicas que son estudiadas por la Minería de Datos y que es necesario profundizar para fortalecer un grupo de investigación destinado a esta ciencia. Una clasificación sumamente interesante y bastante completa es la planteada por Orallo et al [5] en donde se observa en forma más detallada las técnicas particulares que pueden aplicarse según se trate de problemas Predictivos o Descriptivos (Tabla 1).

Nombre	PREDICTIVO		DESCRIPTIVO		
	Clasificación	Regresión	Agrupamiento	Reglas de asociación	Correlaciones / Factorizaciones
Redes neuronales	✓	✓	✓		
Árboles de decisión ID3, C4.5, C5.0	✓				
Árboles de decisión CART	✓	✓			
Otros árboles de decisión	✓	✓	✓	✓	
Redes de Kohonen			✓		
Regresión lineal y logística		✓			✓
Regresión logística	✓			✓	
K-means			✓		
Apriori				✓	
Naive Bayes	✓				
Vecinos más próximos	✓	✓	✓		
Análisis factorial y de comp. ppales.					✓
TwoStep, Cobweb			✓		
Algoritmos genéticos y evolutivos	✓	✓	✓	✓	✓
Máquinas de vectores soporte	✓	✓	✓		
CN2 rules (cobertura)	✓			✓	
Análisis discriminante multivariante	✓				

Tabla 1: Aplicación de técnicas de Data Mining según el tipo de Problema

El trabajo realizado precedentemente por el grupo de investigación se enfocó en la

aplicación de la minería de datos en el ámbito de la seguridad. Sin embargo, se pudo determinar que el conjunto de herramientas que la caracterizan es amplio y surgieron problemas en la integración con los sistemas de donde se extraían la información, la mayoría de ellos sistemas informáticos de plataforma web. Es objetivo del equipo de trabajo investigar el procedimiento de incluir las técnicas de minería de datos en los sistemas que hoy cuentan con la información necesaria para tomar decisiones.

La Inteligencia de negocios o BI, por sus siglas en inglés, según el Data Warehouse Institute, lo define como la combinación de tecnología, herramientas y procesos que permiten transformar los datos almacenados en información, esta información es conocimiento y este conocimiento debe estar dirigido a un plan o tener una estrategia comercial. La inteligencia de negocios debe ser parte de la estrategia empresarial, que permite optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la empresa y la capacidad de tomar buenas decisiones para obtener mejores resultados, tal como puede observar en la imagen1.[8]



Imagen 1: Pasos de la Inteligencia de Negocios

Los pasos que detalla el The Data Warehousing Institute se refiere al proceso de convertir datos en conocimiento, posteriormente reflejar en acciones competitivas. Los datos deben ser procesados

para requerir la información de ella, tarea que las empresas se han visto obligadas a requerir y les permite ser competitivos en el mercado actual. En estas últimas décadas los usuarios consumen productos y servicios por la web y esta modalidad de mercado exige que las empresas deban extraer conocimientos de la información con la que cuentan y puedan predecir acciones. Es sumamente importante la integración de los sistemas con la extracción del conocimiento que se realiza. Se puede afirmar en consecuencia que la etapa de requerir y hacer uso de la información tiene un desafío mayor y consistente en la extracción del conocimiento o análisis inteligente.

Una falencia o una situación a considerar que se presentó en la investigación realizada previamente por este grupo, es la falta de implementación o interrelación de las técnicas empleadas en estas herramientas de Minería de Datos en lenguajes de programación. Existen herramientas en el mercado actual para Minería de Datos, tal como se puede observar en el llamado “Cuadrante Mágico” para plataformas de Análisis Predictivo del año 2.016 (Imagen 2), de la empresa Gartner, organización de investigación de las tecnologías de la información reconocida mundialmente, en donde se muestra la comparación de las principales herramientas para Plataformas Analítica Avanzadas [6], con características y funcionalidades totalmente

diferentes entre ellas.



Imagen 2: “Cuadrante Mágico” para plataformas de Análisis Predictivo 2.016 de Gartner

Hay también, lenguajes de programación que son utilizados mundialmente, donde se suele tomar como medida el Índice “TIOBE Quality Indicator”, el cual es un índice elaborado por una empresa de software holandesa que se especializa en la evaluación y seguimiento de la calidad de los programas informáticos[7], tal como se observa en la imagen 3, revisando actualmente en tiempo real más de 300 millones de códigos de diversos programas.

Aug 2017	Aug 2016	Change	Programming Language	Ratings	Change
1	1		Java	12.961%	-6.65%
2	2		C	6.477%	-4.83%
3	3		C++	6.550%	-0.25%
4	4		OR	4.192%	-0.71%
5	5		Python	3.697%	-0.71%
6	9	▲	Visual Basic .NET	2.369%	+0.03%
7	6	▼	PHP	2.293%	-0.60%
8	7	▼	JavaScript	2.096%	-0.61%
9	8		Perl	1.995%	-0.52%
10	12	▲	Ruby	1.965%	-0.51%

Imagen 3: Comparación de los principales lenguajes, índice TIOBE, agosto de 2.017 De acuerdo a dicho cuadro comparativo se debe hacer hincapié en los lenguajes que son frecuentemente utilizados para desarrollo de sistemas, considerando que en los primeros

lugares se puede visualizar que los lenguajes son de característica web y la integración que se debe desarrollar con la herramientas de la imagen 2, deben estar en esta línea de trabajo. La empresa o entidad que necesita tomar decisiones debe por lo tanto integrar la información que posee y el modelo de extracción del conocimiento, siendo para ello necesario relacionar el modelo de conocimiento con el sistema informático, para que personas de los distintos ámbitos de la empresa puedan realizar el análisis con los datos de la organización.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Dentro de las Líneas Prioritarias de Investigación de la Facultad de Ingeniería (LIPIFI) - UNJu los temas que se investigarán y desarrollarán serian:

- Técnicas Descriptivas de Minería de Datos
- Técnicas Predictivas de Minería de Datos
- Herramientas de Software Libre para Datamining
- Análisis Inteligente de Datos

3. RESULTADOS OBTENIDOS / ESPERADOS

El presente proyecto se encuentra en su fase inicial, y busca cumplir el objetivo general: Implementar técnicas específicas de herramientas de Minería de datos en aplicaciones web de sistemas informáticos con Base de Datos Relacionales; y presenta los siguientes Objetivos específicos:

- Estudiar y profundizar en las principales técnicas de Minería de Datos,
- Analizar y establecer las principales características de las herramientas de Minería de Datos disponibles en la actualidad que permitan su implementación en lenguajes de programación web.
- Estudiar y comparar los principales lenguajes de programación web disponibles en el mercado actual.
- Estudiar e implementar la conectividad de los lenguaje/s de programación web seleccionados con un motor de base de datos relacional.
- Implementar las técnicas de las herramientas seleccionadas en los lenguajes de programación web estudiados.
- Realizar prácticas de Análisis Inteligente en problemas específicos de Minería de Datos.

4. FORMACIÓN DE RECURSOS HUMANOS

Apellido/s y Nombre/s	Formación y Unidad Académica	Rol Dentro del Proyecto
Vargas, Luis Alejandro	Ingeniero en Informática - Fac.Ing.UNJu	Director
Farfán, José Humberto	Esp.Doc.Sup - Ingeniero en Informática - Fac.Ing.UNJu	Co Director
Rodriguez, Mariela Ester	Lic.Sistemas - Ingeniero en	Investigador

	Informática - Fac.Ing.UNJu	
Aramayo, Fernando Ruben	Lic-Sistemas-Ingeniero en Informática - Fac.Cs.Agr. UNJu	Investigador
Ramos, Pablo Nicolás	Estudiante Ingeniería en Informática - Fac.Ing.UNJu	Investigador
Llampa, Alvaro Facundo	Estudiante Ingeniería en Informática - Fac.Ing.UNJu	Investigador
Montes, Leonardo Ezequiel	Estudiante Ingeniería Informática - Fac.Ing.UNJu	Investigador
Mogro, Nelson Ariel	Estudiante Ingeniería en Informática, Fac.Ing.UNJu	Investigador
Córdoba, Irma Rafaela Mercedes	Ingeniera en Informática - Fac.Ing.UNJu	Investigador
Tapia, Marcela Alejandra	Ingeniera en Sistemas de Información Fac.Ing.UNJu	Investigador
Paredes, Julio Cesar	Programador Universitario Fac.Ing.UNJu	Investigador
Spadoni, Gustavo Fernando	Médico Veterinario Fac.Cs.Agr. UNJu	Investigador

Los alumnos Leonardo Ezequiel Montes y Nelson Ariel Mogro se encuentran desarrollando la tesis de grado para la obtención del título de Ingeniero en Informática denominado “Minería de datos para soporte a decisiones de planificación educativa” cuyo Director es el Esp. Ing. José Farfán, su Codirector Ing. Mariela Rodriguez.

5. BIBLIOGRAFIA

- [1] Maimon, O., & Rokac, L., 2010. “Data Mining and Knowledge Discovery Handbook”, “O. Maimon, & L. Rokac, Data Mining and Knowledge Discovery Handbook”, Nueva York, Springer, 2010, págs. 1-18.
- [2] Yelitzza, J., Marcano, A., & Rosalba Talavera, P. , “Minería de Datos como soporte a la toma de decisiones empresariales” , Obtenido de http://www.scielo.org.ve/scielo.php?script=sci_arttext&pid=S1012-15872007000100008 en Agosto de 2.017, Universidad de Zulia, Maracaibo, 2007
- [3] Vallejos, S. J. . “Minería de Datos”, Obtenido de http://exa.unne.edu.ar/informatica/SO/Mineria_Datos_Vallejos.pdf, Universidad Nacional del Nordeste, 2006
- [4] Zhang, S., Zhang, C., & Yang, Q. , “Data preparation for data mining. Applied Artificial Intelligence”, San Francisco, 2003, págs. 375-381.
- [5] Hernández Orallo J., Ramírez Quintana J., Ramirez C.F., “Introducción a la Minería de Datos”, Universidad Politécnica de Valencia, España, Ed.Pearson, 2004, págs 19-42, 137-164, 237-252
- [6] Gartner, “Magic Quadrant for Advanced Analytics Platforms”. Obtenido de <http://www.kdnuggets.com/2017/02/gartner-2017-mq-data-science-platforms-gainers-losers.html> en Agosto de 2.017.
- [7] TIOBE Quality Indicator. The Importance of Being Earnest Quality Indicator. Obtenido de <https://www.tiobe.com/tiobe-index/> en Septiembre de 2.017.
- [8] Oracle, Inteligencia de Negocios. Obtenido de http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529_esa.pdf en setiembre de 2017

VISUALIZACIÓN EN UN ENTORNO DE MINERÍA DE TEXTO

Laura Viviana Gutierrez, Graciela Beguerí, María Alejandra Malberti, Raúl Oscar Klenzi, Manuel Ortega

Departamento de Informática – Facultad de Ciencias Exactas Físicas y Naturales.
Universidad Nacional de San Juan

Ignacio de la Roza y Meglioli. Complejo Universitario Islas Malvinas CUIIM.
Rivadavia, San Juan, Argentina

{gutierrez.laura, grabeda, amalberti, rauloscarklenzi, [manuel.ortega](mailto:manuel.ortega@gmail.com)}@gmail.com

RESUMEN

En el marco de Ciencia de Datos, se propone analizar y caracterizar diferentes estrategias y herramientas de minería de texto, según sus potencialidades de Visualización de Información. Estas se aplicarán a conjuntos de datos obtenidos desde los planes de estudios de las carreras de Informática, disponibles en el Departamento de Informática de la Facultad de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de San Juan DI-FCEFN-UNSJ. Se considerarán herramientas de software libre, particularmente RapidMiner y Knime para la visualización de la información obtenida aplicando técnicas de Textmining.

Considerando al usuario como el destinatario del proceso de búsqueda de conocimiento en datos, se investigará sobre aspectos de interpretación y percepción.

Palabras clave: Visualización, Visualización de la Información, Visualización de Texto, Minería de Texto, Textmining.

CONTEXTO

La línea de investigación se encuentra en el ámbito de los proyectos “Ciencia de los Datos aplicada a grandes colecciones de datos” ejecutado en el bienio 2016_2017 y la continuidad, buscada por el grupo de investigadores que lo conformaban, en la presentación de un proyecto para el bienio 2018_2019 actualmente en evaluación “Visualización y Deep Learning en Ciencia de los Datos”

En este último se pretende evaluar software libre apropiado al área Ciencia de Datos, indagar en la temática de Deep Learning, investigar sobre aspectos de interpretación y percepción relacionados con mecanismos de visualización de información, así como herramientas para implementar los mismos, en el marco de búsqueda de conocimiento en datos. Además de caracterizar a los usuarios de acuerdo a las potencialidades de las herramientas analizadas.

En este contexto, la presente propuesta de trabajo y línea de investigación se centra en formas de visualizar la información generada a partir de los datos de los planes de estudio de las

carreras de informática del DI-FCEF-UNSN, aplicando técnicas de visualización.

1. INTRODUCCIÓN

Según Strecker, Card y otros la Visualización de Información puede ser tratada como un campo de conocimiento bien establecido, asociado al uso de representaciones visuales de datos abstractos que tienen como fin expandir el conocimiento. Para Cairo, A. es una tecnología plural que consiste en transformar datos en información semántica o en la creación de herramientas para dicha transformación, basada en la combinación de señales de naturaleza icónica con otros de naturaleza arbitraria y abstracta (textos, estadísticas, etc.).

La minería de textos es un subcampo de la minería de datos, encontrándose entre sus aplicaciones analizar o comparar textos. Las técnicas de visualización mejoran la comunicación de los resultados obtenidos. En el proceso de análisis se produce la modificación del texto original, también llamado dato no-estructurado, por ejemplo reduciendo un texto a una lista de palabras de acuerdo con su frecuencia. Es así que, la mayoría de las visualizaciones de textos transforman los datos de tipo textual o no estructurados en un nuevo conjunto de datos estructurados, y reducidos, respecto al texto original. Este nuevo conjunto de datos ya no es unidimensional, sino que puede estar ordenado por categorías o con una estructura de red.

En la actualidad se puede contar con varias herramientas libres que ayudan a convertir datos en gráficos. Estas pueden ser usadas desde usuarios principiantes hasta usuarios avezados. En Wikipedia se presentan algunas de

ellas, al igual que en la figura 1, situada en <https://www.bbva.com/wp-content/uploads/2017/10/ebook-cibbva-visualizacion-de-datos-es.pdf>.



Fig.1 – Algunas herramientas para Visualización de datos.

José Mondragón de IBM developerWorks expresa que, en las empresas se generan mucha información estructurada así como datos no estructurados. Opina que combinar los análisis con datamining para ambos tipos de información puede ayudar a incrementar la rentabilidad y la participación en el mercado. En la figura 2 se muestra el diagrama de cómo mejorar el desempeño de un modelo predictivo:

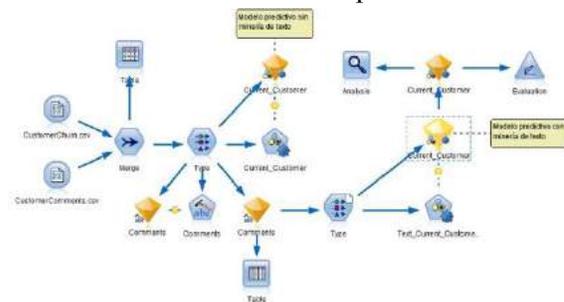


Fig. 2 – Modelo predictivo usando minería de texto.

Es importante tener en cuenta que todo proceso de minería de texto debería culminar con la confección de un

gráfico mediante el cual se visualice la información de manera rápida y clara, tal como lo indica Felipe de Jesús Núñez Cárdena

El punto de partida en el tratamiento de visualizaciones de textos, es reconocer dos grandes categorías de textos: Textos individuales y Colecciones de textos, y a la vez consideran distintos tratamientos de visualizaciones relacionados con cada una de las categorías.

1 - Visualización de textos individuales

Los métodos de visualización de un texto completo suelen utilizar el color como elemento distintivo, variaciones de diagramas de barras, curvas que conectan partes del texto entre otros recursos. Una problemática detectada es la variedad en cuanto a la naturaleza de los textos, por lo que los autores mencionados consideran que es una buena idea trabajar con textos más estructurados y sencillos, con un vocabulario más regular, una longitud del texto estandarizada, con una clara estructura del discurso y corrección en el lenguaje (artículos científicos, textos de patentes, diagnósticos de salud, etc.).

Si la problemática de visualización se traslada a partes de un texto, un método habitual es el llamado bag of words o bolsa de palabras

Asimismo, métodos estadísticos muy simples, como la frecuencia de palabras, pueden tener un resultado fácil de entender. También una lista de varios tamaños de palabras es una forma directa de comunicarse con cualquier usuario, ya sean estos principiantes o expertos.

2 - Visualización de colecciones

Una vez determinadas las colecciones, los datos se pueden considerar como un caso general de visualización de datos. Es así que se emplean métodos utilizados en otros campos como ser visualizaciones de red, líneas temporales, ítems de nubes de palabras

-para el caso de comparar textos, y el color para detectar palabras nuevas (Jaume Nualart-Vilaplana, Mario Pérez-Montoro y Mitchell Whitelaw).

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

En el marco del proyecto que contiene la presente línea de investigación se pretende, tomar el trabajo realizado en la tesis de maestría de la Lic. Laura Gutiérrez. Este versa sobre proponer una metodología automática para determinar las pertinencias sintáctico-semánticas entre los contenidos mínimos de carreras de informática con las normas establecidas según las diferentes titulaciones asignadas a las mismas.

El proceso automático agrupa los contenidos mínimos de las carreras estableciendo el grado de similitud sintáctica de estos, con las áreas u objetivos temáticos de las resoluciones o marcos regulatorios correspondientes, favoreciendo así, un análisis inicial de planes de estudio que pueda llevar a la decisión de hacer correcciones sobre los mismos.

La aplicación se llevó adelante mediante la utilización los módulos de modelado y minería de texto (TextMining -TM-) de la herramienta de software libre RapidMiner (RM) versión 5.3.15. El caso de estudio considerado es la carrera Licenciatura en Ciencias de la Computación del DIFCFN-UNSJ.

Ese trabajo se ha llevado adelante mediante la utilización de herramientas de software libre del área de Data mining - DM. En este caso los resultados que se presentaron se alcanzaron mediante el uso de algoritmos de DM que posee la herramienta RapidMiner versión 5.3.15 con la opción de visualizarlos también.

Desarrollo

El trabajo citado se realizó básicamente en tres pasos:

1_ Preprocesamiento y Análisis de Documentos de Texto (Planes de Estudio)

2_ **Visualización de Contenidos Faltantes:** El resultado del análisis a los datos, permitió obtener los contenidos faltantes en los documentos de texto utilizados, que con los documentos de texto a simple vista no se observaban (Contenidos mínimos).

3_ Determinación de pertinencias sintácticas (Planes de Estudios-Ordenanzas)

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos en este trabajo en particular, permitieron visualizar los contenidos faltantes en los documentos de texto y llevaron a sugerir cambios en los mismos obteniendo mejores resultados en un posterior análisis.

Se pretende desde este proyecto aplicar distintas herramientas de visualización en diferentes conjuntos de datos textuales, con el propósito de extraer rápida y sencillamente el conocimiento.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación se encuentra conformado por: una directora, un co-director, cuatro docentes investigadores (categorizados en el Programa de Incentivos de la Secretaria de Políticas Universitarias (SPU) perteneciente al

Ministerio de Educación de la Nación Argentina), un egresado y cuatro alumnos de los últimos años de las carreras del Departamento de Informática.

En el marco de esta investigación se desarrollarán trabajos finales para la/s carrera/s LCC-LSI del DI.

La ejecución del proyecto incidirá directamente en una formación más profunda de los integrantes del equipo de investigación. Este aspecto beneficiará de manera directa a las carreras del Departamento de Informática, pues las temáticas abordadas están vinculadas con las materias en las cuales se desempeñan los integrantes de este proyecto. También estas líneas de trabajo servirán al medio para proveer nuevas estrategias de administración y presentación del conocimiento.

5. BIBLIOGRAFÍA

- Cairo, A. El arte funcional: infografía y visualización de información. Alamut, 2011. ISBN:9788498890679.
- Card, S. K.; Mackinlay, J. D.; Shneiderman, B. Readings in information visualization: using vision to think. Morgan Kaufmann, 1999. ISBN-10:1558605339.
- De Lucia Castillo, F., & Saibel Santos, C. A. (2016). Nubes de palabras animadas para la visualización de información textual de Publicaciones Académicas.
- Dürsteler, J. C. (2000). Visualización de información. *Gestion*.
- El Cairo, A. (2012). *El arte funcional: una introducción a*

gráficos de información y visualización. Nuevos jinetes.

- Gutiérrez, L., Klenzi, R., Malberti, A., Beguerí, G., Pinto, T., & Araya, J. (2015). Análisis de Planes de Estudio Mediante Determinación Automática de Pertinencias Sintáctico-Temáticas en Carreras de Informática. *Revista Eletrônica Argentina-Brasil De Tecnologias Da InformaçãO E Da ComunicaçãO*, 1(2). doi:10.5281/zenodo.59453
- Hernández Orallo J., Ramirez Quintana, J, Ferri Ramirez, C. (2008) *Introducción a la Minería de Datos*. Pearson-Prentice Hall.
- https://es.wikipedia.org/wiki/Visualizaci%C3%B3n_de_datos
- Kargupta, H., Han, J., Philip, S. Y., Motwani, R., & Kumar, V. (Eds.). Next generation of data mining. CRC Press. (2008).
- Larose, D. (2006) *Data Mining. Methods and Models*. Department of Mathematical Sciences Central Connecticut State University Wiley. A John Wiley & Sons, Inc Publication.
- Larose, D. (2005) *Discovering Knowledge In Data - An Introduction to Data Mining*. John Wiley & Sons, Inc., Publication.
- Miller, T. W. (2005). *Data and text mining: A business applications approach* (pp. 917-2199). New Jersey: Pearson Prentice Hall.
- Nualart-Vilaplana, J., Pérez-Montoro, M., & Whitelaw, M. (2014). Cómo dibujamos textos: Revisión de propuestas de visualización y exploración textual. *El profesional de la información*, 23(3), 221-235.
- Strecker, J. Data visualization in review:summary. Tech. Report IDRC, 2012. URL: <http://idlbnc.idrc.ca/dspace/bitstream/10625/49286/1/IDL-49286.pdf>

PROPUESTA DE PROCEDIMIENTO PARA EL ANÁLISIS DELICTIVO BASADO EN LA EXPLOTACIÓN DE LA INFORMACIÓN

Lorena Flores¹, Sonia Mariño², Sebastian Martins³

¹Maestría en Tecnologías de la Información, Universidad Nacional del Nordeste. Corrientes. Argentina lorenaelizabeth.flores@gmail.com

²Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura. 9 de Julio 1449. CP: 3400. Corrientes. Argentina. Universidad Nacional del Nordeste, simarinio@yahoo.com

³Laboratorio de Investigación y Desarrollo en Ingeniería de Explotación de Información Grupo Investigación en Sistemas de Información. Departamento de Desarrollo Productivo y Tecnológico. Remedios de Escalada, Lanús. Argentina. Universidad Nacional de Lanús smartins089@gmail.com

CONTEXTO

El proyecto se desarrolla en el marco del proyecto “TI en los sistemas de información: modelos, métodos y herramientas” acreditado por la Universidad Nacional del Nordeste (UNNE) y la Maestría en Tecnologías de la Información (UNNE-UNAM).

RESUMEN

Los proyectos de explotación de información son ampliamente desarrollados en diversos dominios de conocimiento. En la sociedad actual se aplican diversas tecnologías de apoyo a la gestión de incidentes. Por ello, se propone diseñar un procedimiento para la toma de decisiones integrando tecnología GIS y técnicas de minería de datos. La propuesta se validará la propuesta en un caso de estudio circunscripto a la Provincia de Corrientes. Se espera transferir los conocimientos adquiridos y los productos generados a instituciones públicas provinciales.

Palabras clave: Minería de datos, Explotación de información, Modelo, GIS, incidentes delictivos

1. INTRODUCCIÓN

Un sistema de información geográfica (GIS) se define como un sistema informático que

permite capturar, almacenar, consultar, analizar y mostrar datos geoespaciales [1]. Estos sistemas son ampliamente utilizados en diversos dominios atendiendo a las innovadoras alternativas de gestión de la información que ofrecen.

La tecnología GIS se utiliza en el tratamiento de la información de diversos dominios, siendo uno de ellos el referente a la delincuencia. En [2] se exponen herramientas para el mapeo efectivo del crimen, análisis y gestión. En este sentido se coincide con lo presentado en [3], dado que actualmente las evidencias digitales para la investigación del crimen en su mayoría incluyen un componente geoespacial.

Dado que los GIS pueden tratarse como bases de datos que contienen información geográfica de utilidad vinculada a un determinado territorio de interés. Desde esta visión, pueden aplicarse para la búsqueda de patrones y regularidades significativas utilizando ingeniería de explotación de información [4].

La explotación de información (information mining) constituye la sub-disciplina de la Informática que aporta a la Inteligencia de Negocio [5], métodos y herramientas para la transformación de información en conocimiento [4]. Un proceso de explotación de información se puede definir como un conjunto de tareas relacionadas lógicamente

[6] que se ejecutan para lograr, a partir de un conjunto de información con un grado de valor para la organización, otro conjunto de información con mayor grado de valor que el inicial [7, 8].

Para lograr este objetivo se utilizan las técnicas de minería de datos (data mining o MD). La minería de datos es la extracción de información no trivial, implícita, previamente desconocida y potencialmente útil de una base de datos [9]. Es un elemento fundamental para la explotación de información o bien del proceso que tiene como objetivo el descubrimiento de conocimiento en grandes bases de datos (knowledge discovery in databases o KDD) [10, 11].

Los procesos de explotación de información basados en sistemas inteligentes [12], se centran en el descubrimiento de patrones de conocimiento en la masa de datos, aplicando técnicas de minería de datos. Sin embargo se carece de una visión o enfoque de cómo extender estas técnicas de descubrimiento de patrones en la información contenida en una base de datos geo-referenciada generada y utilizada por un GIS. Lo expuesto permitirá comprender la distribución espacial de una temática en relación a un contexto geográfico específico.

En [13, 15] se reportan análisis de incidentes vinculados con la criminalidad.

Una de las medidas para combatir la delincuencia en el país fue la creación del Sistema de Alerta Temprana (SAT) por parte del Ministerio de Justicia y Derechos Humanos [13]. En la actualidad, la información proveniente del SAT se trata a través del análisis estadístico, sin aplicar técnicas ni herramientas de minería de datos. Además, en los últimos años, el Instituto Tecnológico de Buenos Aires (ITBA) en conjunto con otras Universidades, realiza estudios enfocados en el campo delictivo donde da cuenta de la importancia de aplicar

la minería de datos para la exploración y detección de patrones delictivos en Argentina [16], como así también en la inteligencia criminal [17].

En cuanto a métodos de integración de tecnologías GIS y MD, se presentan propuestas para integrar ambas herramientas, en los cuales la mayoría de ellos solo se limita a mostrar un ejemplo o caso de uso para ilustrar la viabilidad [21, 22].

El estudio demuestra que es sumamente necesaria la validación de las propuestas para obtener resultados empíricos sobre su efectividad. Del mismo modo, se puede mencionar un trabajo de la Universidad Nacional de Lanús en el marco del cual se desarrolló una extensión de software GIS para integrar minería de datos [23].

Otros relevamientos y estudios realizados en la región del NEA indican la carencia de la aplicación de estos métodos y herramientas software. Así, se nota el aporte posible a los departamentos policiales al integrar la aplicación de las herramientas GIS y las técnicas de minería de datos como instrumentos para reducir la criminalidad.

Por lo expuesto, se propone integrar los sistemas de información geográfica y las técnicas de minería de datos como herramientas claves para la detención y predicción de hechos delictivos.

Para validar la propuesta, se aplicarán estas tecnologías sobre una base de datos geo-referenciada con información de ocurrencias de delitos de robo y hurto cometidos en la ciudad en el primer semestre del año 2017, con el propósito de realizar un análisis y hallazgo de los patrones más relevantes de estos hechos, y así aportar una innovadora herramienta estratégica y de toma de decisiones.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el marco del trabajo planteado se indagará en:

- Los casos reportados para explotar información capturada por tecnologías GIS y aplicando métodos de minería de datos.
- Las metodologías descritas en la literatura para ejecutar proyectos de explotación de información.
- Las herramientas Open Source GIS para la captura de información.
- Los procesos de explotación de la información aplicables para el procesamiento de datos delictivos capturados por tecnologías GIS.
- Las herramientas Open Source para aplicar procesos de minería de datos que posibilitan el desarrollo de una solución tecnológica.

3. RESULTADOS OBTENIDOS/ESPERADOS

El proyecto que se expone tiene como objetivo general aportar un procedimiento que integre las tecnologías GIS y los métodos comprendidos en la minería de datos para identificar patrones delictivos en un contexto geográfico delimitado. Este objetivo se concretará a través de:

- Estudiar, analizar y seleccionar tecnologías de geo-referenciación GIS/IDE existentes.
- Investigar y profundizar en técnicas y herramientas comprendidas en la minería de datos.
- Elaborar y desarrollar una propuesta integradora de tecnologías GIS y métodos comprendidos en la minería de datos, para aplicar procesos de explotación de la información en un dominio específico
- Validar la propuesta, en un caso de estudio, para producir conocimiento

orientado a la detección de hechos delictivos de robos y hurtos en la ciudad de Corrientes.

El proyecto se inició en diciembre 2017, como resultados obtenidos se menciona la elaboración de la Revisión Sistemática de la Literatura referente a la integración de MD y GIS aplicados al análisis delictivo. El informe elaborado aportó en la delimitación del objeto de estudio y en la definición del plan de trabajo presentado a la asignatura Trabajo Final de Maestría.

Se espera que el desarrollo de la propuesta contribuya a: sistematizar las tareas manuales de los agentes policiales y a la generación de información de calidad que apoye la toma de decisiones. Esta se sustentará en los métodos de explotación de la información aplicados a los datos capturados y procesados con tecnologías GIS por el SAT.

4. FORMACIÓN DE RECURSOS HUMANOS

Desde el proyecto de I+D se fortalece la formación de recursos humanos en temas que aportan al desarrollo tecnológico regional. Además, el trabajo forma parte de un proyecto de trabajo final de la Maestría en Tecnologías de la Información, carrera cooperativa entre la UNNE y la UNAM.

El conocimiento adquirido y consolidado a través de la propuesta se podrá transferir a los recursos humanos de la organización adoptante de la tecnología, y la información producida a la ciudadanía. En el futuro se replicará el procedimiento en otros dominios del conocimiento.

5. BIBLIOGRAFIA

- [1] K. T. Chang. Introduction to Geographic Information Systems. 7th Ed. New York: McGraw-Hill, 2014
- [2] E.M.A. El-Aziz, S. Mesbah and K. Mahar, "GIS-Based Decision Support System for Criminal

- Tracking”, in 22nd International Conference on Computer Theory and Applications (ICCTA), Alexandria, Egypt, 2012, pp. 31.
- [3] A. Tillekens, N.A. Le-Khac and T.T.P. Thi, “A Bespoke Forensics GIS Tool”, in International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2016, pp. 1-2.
- [4] R. García Martínez, P. Britos, P. Pesado, R. Bertone, F. Pollo-Cattaneo, D. Rodríguez, P. Pytel and J. Vanrell, “Towards an Information Mining Engineering,” in Software Engineering, Methods, Modeling and Teaching. Sello Editorial Universidad de Medellín. ISBN 978-958-8692-32-6. pp. 83-99.2011.
- [5] S. Negash and P. Gray, “Business Intelligence,” In Handbook on Decision Support Systems. 2ds. Ed. (F. Burstein and C. Holsapple, Springer), 2008, pp. 175-193.
- [6] B. Curtis, M. Kellner and J. Over, “Process Modelling”, Communications of the ACM, vol. 35, no. 9, pp. 75-90, September 1992.
- [7] J. Ferreira, O. K. Takai and C. Pu, “Integration of business processes with autonomous information systems: a case study in government services”, in IEEE Conference on Electronic Commerce Technology. Munich, Germany, 2005, pp. 471–474.
- [8] I. Hann, K. Hui, S. Lee and I. Png, “Analyzing Online Information Privacy Concerns: An Information Processing Theory Approach”, in Proceedings 40th Annual Hawaii International Conference on System Sciences. Waikoloa, HI, USA, 2007, pp. 210-219.
- [9] W. Frawley, G. Piatetsky-Shapiro and C. Matheus, “Knowledge Discovery in Databases: An Overview”, AI Magazine, vol 13, no 5, pp. 58, 1992.
- [10] U. Fayyad, G. Piatetsky-Shapiro and P. Smyth, “From Data Mining to Knowledge Discovery in Databases”, AI Magazine, vol. 17, no. 3, pp. 39, 1996.
- [11] R. García Martínez, “Sistemas Autónomos,” en Aprendizaje Automático. Editorial Nueva Librería. ISBN 950-9088-84-6, 1997.
- [12] P. Britos, “Procesos de explotación de información basados en sistemas inteligentes,” Tesis Doctoral en Ciencias Informáticas, Facultad de Informática. Universidad Nacional de la Plata. Buenos Aires, Argentina, 2008.
- [13] I. Perversi, F. Valenga, E. Fernandez, P. Britos, and R. Garcia-Martinez, “Identificación y detección de patrones delictivos basada en minería de datos”, IX Workshop de Investigadores en Ciencias de la Computación. Argentina, 2007, pp. 1-5.
- [15] H. Chen, W. Chung, J. J. Xu, G. Wang, Y. Qin and M. Chau, “Crime Data Mining: A General Framework and Some Examples” Computer, vol 37, no 4, April, 2004.
- [16] F. Valenga, I. Perversi, E. Fernández, H. Merlino, D. Rodríguez, P. Britos y R. García-Martínez, “Aplicación de minería de datos para la exploración y detección de patrones delictivos en Argentina”, en XIII Congreso Argentino de Ciencias de la Computación. Argentina, 2007.
- [17] P. Britos, E. Fernández, H. Merlino, F. Pollo-Cataneo, D. Rodríguez, C. Procopio, C. Rancan y R. García-Martínez, “Explotación de información aplicada a inteligencia

- criminal en Argentina”, en XIII Congreso Argentino de Ciencias de la Computación. Argentina, 2008.
- [21] C. J. Chang and S. W. Shyue, “Association Rules Mining with GIS: An Application to Taiwan Census 2000”, in Sixth International Conference on Fuzzy Systems and Knowledge Discovery. Tianjin, China. 2009.
- [22] J. Wang, X. Chen, K. Zhou, H. Zhang and W. Wang, “Research of GIS-based spatial data mining model”, in Second International Workshop on Knowledge Discovery and Data Mining, Moscow, Russia. 2009.
- [23] A. Segura, P. Santamaria, F. Mieres y D. Aguirre, “Ambiente de Explotación de Información para Sistemas de Información Geográfica”, Departamento de Desarrollo Productivo y Tecnológico, Universidad Nacional de Lanús, Buenos Aires, Argentina. 2003. [En línea] Disponible: <http://sistemas.unla.edu.ar/sistemas/sls/ls-3-Proyecto-de-Software/pdf/2013-TF-PdS-EI-SIG.pdf> [Accedido: 19-Diciembre-2017]

Indexación y Búsqueda sobre Datos no Estructurados

Norma Herrera, Darío Ruano, Paola Azar

Departamento de Informática

Universidad Nacional de San Luis, Argentina

{nherrera, dmruano, epazar}@unsl.edu.ar

Anabella De Battista, Andrés Pascal

Departamento Ingeniería en Sistemas de Información

FRCU, Universidad Tecnológica Nacional, Entre Ríos, Argentina

{anadebattista, andrespascal22}@gmail.com

Abstract

Las bases de datos han incluido la capacidad de almacenar datos no estructurados tales como imágenes, sonido, texto, video, etc. La problemática de almacenamiento y búsqueda en estos tipos de base de datos difiere de las bases de datos clásicas, dado que no es posible organizarlos en registros y campos, y aun cuando pudiera hacerse, la búsqueda exacta carece de interés. Es en este contexto donde surgen nuevos modelos de bases de datos capaces de cubrir las necesidades de almacenamiento y búsqueda de estas aplicaciones. Nuestro interés se basa en el diseño de índices eficientes para estas nuevas bases de datos.

1 Contexto

El presente trabajo se desarrolla en el ámbito de la línea Técnicas de Indexación para Datos no Estructurados del Proyecto Tecnologías Avanzadas de Bases de Datos (22/F414), cuyo objetivo es realizar investigación básica sobre manejo y recuperación eficiente de información no tradicional.

2 Introducción

La información disponible en formato digital aumenta día a día su tamaño de manera exponencial. Gran parte de esta información involucra el uso de datos no estructurados tales como imágenes, sonido, texto, video, etc. Debido a que no es posible organizar estos tipos de datos en registros y campos, las tecnologías tradicionales de bases de datos para almacenamiento y búsqueda de información no son adecuadas en este ámbito.

Es en este contexto donde surgen nuevos modelos de bases de datos capaces de cubrir las necesidades de almacenamiento y búsqueda de estas aplicaciones. Nuestro interés se basa en el diseño de índices eficientes para estas nuevas bases de datos.

Bases de Datos Textuales (BDT) Una base de datos de texto es un sistema que mantiene una colección grande de texto, y provee acceso rápido y seguro al mismo. Sin pérdida de generalidad, asumiremos que la base de datos de texto es un único texto T posiblemente almacenado en varios archivos. Las búsquedas en la que el usuario ingresa un *patrón de búsqueda*

y el sistema retorna todas las posiciones del texto donde el patrón ocurre, es una de las búsquedas más comunes en este tipo de bases de datos.

Mientras que en bases de datos tradicionales los índices ocupan menos espacio que el conjunto de datos indexado, en las bases de datos de texto el índice ocupa más espacio que el texto, pudiendo necesitar de 4 a 20 veces el tamaño del mismo [5, 8]. Una alternativa para reducir el espacio ocupado por el índice es buscar una representación compacta del mismo, manteniendo las facilidades de navegación sobre la estructura. Pero en grandes colecciones de texto, el índice aún comprimido suele ser demasiado grande como para residir en memoria principal [6, 7]. Por esta razón, el estudio de índices comprimidos y en memoria secundaria para búsquedas en texto es un tema de creciente interés.

Espacios Métricos El modelo de espacios métricos permite formalizar el concepto de búsqueda por similitud en bases de datos no tradicionales [2]. Un espacio métrico está formado por un conjunto de objetos \mathcal{X} y una función de distancia d definida entre ellos que mide cuan diferentes son. La base de datos será un subconjunto finito $\mathcal{U} \subseteq \mathcal{X}$.

Una de las consultas más comunes en este modelo de bases de datos es la *búsqueda por rango*. En esta búsqueda dado un elemento $q \in \mathcal{X}$, al que llamaremos *query* y un radio de tolerancia r , la búsqueda por rango consiste en recuperar los objetos de la base de datos cuya distancia a q no sea mayor que r . Para evitar examinar exhaustivamente la base de datos, se preprocesa la misma por medio de un *algoritmo de indexación* con el objetivo de construir una *índice*, diseñado para ahorrar cálculos en el momento de la búsqueda. En [2] se presenta un desarrollo unificador de las soluciones existentes en la temática.

Bases de datos temporales

Las base de datos temporales permiten almacenar y recuperar datos que dependen del tiempo. Mientras que las bases de datos tradicionales tratan al tiempo como otro tipo de dato más, este tipo de base de datos incorpora al tiempo como una dimensión, distinguiendo dos tipos de tiempos:

Tiempo válido expresa el tiempo durante el cual una proposición es cierta.

Tiempo transaccional: indica el momento que los datos fueron incorporados a la base de datos.

Bases de datos métrico-temporales (BDMT)

Este modelo permite almacenar objetos no estructurados con tiempos de vigencia asociados y realizar consultas por similitud y por tiempo en forma simultánea. Formalmente un *Espacio Métrico-Temporal* es un par (U, d) , donde $U = O \times N \times N$, y la función d es de la forma $d : O \times O \rightarrow R^+$. Cada elemento $u \in U$ es una tripla (obj, t_i, t_f) , donde obj es un objeto (por ejemplo, una imagen, sonido, cadena, etc) y $[t_i, t_f]$ es el intervalo de vigencia de obj . La función de distancia d , que mide la similitud entre dos objetos, cumple con las propiedades de una métrica (positividad, simetría y desigualdad triangular).

Un nuevo tipo de consulta son las denominadas métrico-temporales que se definen formalmente en símbolos como:

$$(q, r, t_{iq}, t_{fq})_d = \{o / (o, t_{io}, t_{fo}) \in X \wedge d(q, o) \leq r \wedge (t_{io} \leq t_{fq}) \wedge (t_{iq} \leq t_{fo})\}$$

La consulta implica buscar todos los objetos o de la parte finita X del universo U que estén a una distancia a lo más r de q , y que su tiempo asociado t coincida (o se solape) con en tiempo de la consulta.

Varios índices métrico-temporales se han propuesto en este ámbito, todos estos índices fueron desarrollados para ser eficientes en memoria principal.

3 Líneas de Investigación

3.1 Memoria Secundaria

3.1.1 Base de Datos Textuales

Un *trie de sufijos* es un índice que permite resolver eficientemente las operaciones de búsquedas en texto pero que necesita en espacio 10 veces el tamaño del texto indexado.

En [10] se presenta una nueva representación de un trie de sufijos que permite reducir el espacio necesario para almacenar el índice, eliminando la necesidad de mantener los punteros explícitos a los hijos. Esta representación tiene la ventaja de permitir un posterior proceso de paginado para manejar eficientemente el trie de sufijos en memoria secundaria [11].

Hemos diseñado una técnica de paginación para este índice basándonos en la representación compacta del mismo. Nuestra técnica de paginación consiste en una extensión para árboles r-arios del paginado utilizado en el Compact Pat Tree [3].

Al igual que el algoritmo presentado en [3], nuestra técnica de paginado también particiona el árbol en componentes conexas, denominadas *partes*, cada una de las cuales se almacena en una página de disco.

El algoritmo procede en forma bottom-up tratando de condensar en una única parte un nodo con uno o más de los subárboles que dependen de él. En este proceso de particionado las decisiones se toman en base a la profundidad de cada nodo involucrado, donde la profundidad indica la cantidad de accesos a disco que deberá realizar el proceso de búsqueda para llegar desde esa parte a una hoja del árbol.

Para particionar un árbol, el algoritmo comienza asignando cada hoja a una parte con profundidad 1 y luego, en forma bottom-up, procesa cada uno de los nodos de este árbol r-ario según las reglas que se explican a continuación.

Sea x el nodo corriente a procesar. Se ordenan los hijos del x de mayor a menor según su profundidad, y para aquellos hijos de igual profundidad se ordenan de menor a mayor según su tamaño. Se pueden presentar los siguientes casos:

Caso 1: x y su primer hijo de mayor profundidad d entran en una página de disco:

- Colocamos en una misma parte x y tantos hijos de x como entren en una página, teniendo en cuenta en este proceso el orden ya establecido.
- Si en algún momento nos encontramos que un hijo de x tiene profundidad j y no entra en la página, los siguientes hijos con profundidad j tampoco entrarán (porque están ordenados por tamaños). En este caso se prosigue con los hijos con profundidad menor que j .
- Aplicamos el mismo proceso hasta recorrer todos los grupos de distintas profundidades que existan para los hijos de x .
- Se cierran las partes de aquellos hijos que no conforman la nueva parte creada.
- Si todos los hijos de mayor profundidad d se han agregado a la nueva parte creada, se establece que esta nueva parte tiene profundidad d .
- Si algún hijo de mayor profundidad d es cerrado, la profundidad de la nueva parte se establece en $d + 1$.

Caso 2: x y su primer hijo de mayor profundidad d no entran en una página de disco. En este caso se cierran todas las partes hijas y se crea una nueva parte para el nodo corriente con profundidad $d+1$, donde d es el máximo de las profundidades de los hijos.

Estamos finalizando la implementación de esta propuesta para ya dar inicio a la evaluación experimental.

3.1.2 Base de Datos Métrico-Temporal

Varios índices métrico-temporales se han propuesto en este ámbito, todos estos índices fueron desarrollados para ser eficientes en memoria principal. Uno de ellos es el H-FHQT presentado en [4], que resulta competitivo en aquellos casos donde los objetos tienen vigencia en un sólo instante de tiempo. Nuestro objetivo es proponer una técnica de paginación para que este índice resulte eficiente en memoria secundaria.

El H-FHQT consiste en una lista de los instantes válidos de tiempo, donde cada celda de la lista contiene un índice FHQT [1] con el que indexa todos los objetos vigentes en dicho instante. Nuestra técnica de paginación tiene los siguientes casos a considerar:

Caso 1: La lista de instantes de tiempos válidos entra en memoria primaria pero cada árbol correspondiente a cada instante de tiempo reside en memoria secundaria. En este caso hay dos situaciones a tener en cuenta:

- Cada árbol FHQT correspondiente a cada instante de tiempo entra en una página de disco. En este caso la paginación es directa, haciendo corresponder cada FHQT con una página de disco.
- Cada árbol FHQT correspondiente a cada instante de tiempo no entra en una página de disco, en este caso se procede a paginarlo con la técnica propuesta en la sección anterior.

Caso 2: Ni la lista de instantes de tiempos válidos ni cada uno de los árboles FHQT correspondientes a cada instante de tiempo entran en memoria principal. En este caso utilizamos para la lista de instantes de tiempo válido un árbol B (dado que allí se buscará por igualdad) y para los árboles FHQT usamos la técnica de paginado explicada en la sección anterior.

Cabe señalar que no existe hasta el momento ningún índice en memoria secundaria para este tipo de base de datos.

3.2 Índices Comprimidos

El objetivo de esta línea es el estudio de índices comprimidos para base de datos de texto con el fin de proponer técnicas alternativas que permitan una mejora en tiempo del mismo, pero que siga siendo competitivo en espacio. Como ya lo mencionáramos en la sección anterior, se ha realizado un estudio detallado de la representación compacta del trie de sufijos [10] y de los algoritmos de búsqueda sobre esta representación.

En [10] se propone mantener la topología del árbol a través de la representación de paréntesis [9]. En esta representación, para poder navegar eficientemente el árbol resolver las siguientes operaciones sobre secuencias binarias:

findclose(i) : retorna la posición del paréntesis que cierra que corresponde al paréntesis que abre en la posición i .

findopen(i) : retorna la posición del paréntesis que abre que corresponde al paréntesis que cierra en la posición i .

excess(i) : retorna la diferencia del número de paréntesis que abren y el número de paréntesis que cierra desde el principio a la posición i .

enclose(i) : retorna la posición del paréntesis más cercano que encierra al nodo de la posición i .

Empíricamente hemos detectado que una de las funciones más costosas de ejecutar es *findclose*(i) por lo que desarrollamos una técnica para mejorar el desempeño de esta función, y con ello el desempeño de toda la estructura. La técnica diseñada consiste en obtener información en ciertos niveles del trie de sufijos para que esta función se realice con mayor eficiencia en esos niveles. El nivel hasta el

cual se debe realizar la optimización es un parámetro de nuestro algoritmo que se establecerá empíricamente. Ya se ha terminado la etapa de desarrollo y nos encontramos en la etapa evaluación experimental.

4 Resultados Esperados

Se espera obtener índices eficientes, tanto en espacio como en tiempo, para el procesamiento de consultas en bases de datos textuales y en espacios métricos. Los mismos serán evaluados tanto analíticamente como empíricamente.

5 Recursos Humanos

El trabajo desarrollado en esta línea forma parte del desarrollo de un Trabajo Final de la Licenciatura, dos Tesis de Maestría y una Tesis de Doctorado, todas ellas en el área temática de Ciencias de la Computación, en la Universidad Nacional de San Luis.

References

- [1] R. Baeza-Yates, W. Cunto, U. Manber, and S. Wu. Proximity matching using fixed-queries trees. In *Proc. 5th Combinatorial Pattern Matching (CPM'94)*, LNCS 807, pages 198–212, 1994.
- [2] E. Chávez, G. Navarro, R. Baeza-Yates, and J.L. Marroquín. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273–321, September 2001.
- [3] D. Clark and I. Munro. Efficient suffix tree on secondary storage. In *Proc. 7th ACM-SIAM Symposium on Discrete Algorithms*, pages 383–391, 1996.
- [4] A. De Battista, A. Pascal, G. Gutierrez, and N. Herrera. Un nuevo índice métrico-temporal: el historical-fhqt. In *Actas del XIII Congreso Argentino de Ciencias de la Computación*, Corrientes, Argentina, 2007.
- [5] G. H. Gonnet, R. Baeza-Yates, and T. Snider. *New indices for text: PAT trees and PAT arrays*, pages 66–82. Prentice Hall, New Jersey, 1992.
- [6] R. González and G. Navarro. A compressed text index on secondary memory. In *Proc. 18th International Workshop on Combinatorial Algorithms (IWOCA)*, pages 80–91. College Publications, UK, 2007.
- [7] R. González and G. Navarro. Compressed text indexes with fast locate. In *Proc. 18th Annual Symposium on Combinatorial Pattern Matching (CPM)*, LNCS 4580, pages 216–227, 2007.
- [8] U. Manber and G. Myers. Suffix arrays: A new method for on-line string searches. *SIAM Journal of Computing*, 22(5):935–948, 1993.
- [9] J. Ian Munro and Venkatesh Raman. Succinct representation of balanced parentheses and static trees. *SIAM J. Comput.*, 31(3):762–776, 2001.
- [10] D. Ruano and N. Herrera. Representación secuencial de un trie de sufijos. In *XX Congreso Argentino de Ciencias de la Computación*, Buenos Aires, Argentina, 2014.
- [11] J. Vitter. External memory algorithms and data structures: Dealing with massive data. *ACM Computing Surveys*, 33(2):209–271, 2001.

Algoritmos de aprendizaje automático para respuestas en tiempo real sobre entornos masivos de datos

Banchero Santiago, Tolosa Gabriel H., Tonin Monzón Francisco,
Fernandez Juan M., Paz Soldan Carlos, Giordano Luis A.,
Marrone Agustín H.

{sbanchero, tolosoft, ftonin, jmfernandez, cpazsoldan, agiordano, amarrone}@unlu.edu.ar
Universidad Nacional de Luján

Resumen

En la actualidad existen incontables fuentes de información en tiempo real que provienen de redes de sensores, plataformas de observación del tiempo, mediciones de gases, observación de la tierra desde plataformas satelitales, ciudades inteligentes, entre un sin número de instrumentos que censan y transmiten datos. A su vez hay una creciente demanda por el desarrollo de herramientas para poder extraer conocimiento a partir de esos grandes repositorios de datos. El aprendizaje automático es un área de la inteligencia artificial donde sus métodos contribuyen en el proceso de descubrimiento de conocimiento para la toma de decisiones inteligentes. Las demandas para la extracción de conocimiento en entornos de Big Data ha acrecentado el interés por la utilización de técnicas tradicionales de aprendizaje automático en distintos problemas de repositorios masivos de datos y también en entornos de flujos (o *streaming*) de datos donde muchas veces no es posible su almacenamiento, pero se requiere tomar decisiones en tiempo real conforme se leen los datos.

Contexto

Este proyecto es el comienzo de una nueva línea de investigación del Departamento de Ciencias Básicas (UNLu) que tiene como principal aspira-

ción profundizar los conocimientos sobre métodos actuales de aprendizaje de máquina como herramienta para el descubrimiento de conocimiento en problemas de Big Data sobre *streaming* de datos de diversas naturaleza.

Introducción

En la actualidad existen muchas aplicaciones que hacen un uso intensivo de datos, haciendo que el volumen y la complejidad de estos crezcan rápidamente. Los motores de búsqueda, redes sociales, e-ciencia (por ejemplo: genómica, meteorología y salud), financieros (por ejemplo: banca y megatiendas entre otros) son algunos ejemplos de tales aplicaciones [1, 11]. Esta problemática se conoce como el problema de Big Data [13].

Big Data se caracteriza por tres aspectos: (a) los datos son numerosos, (b) los datos no pueden ser categorizados en bases de datos relacionales regulares, y (c) los datos son generados, capturados y procesados muy rápidamente [10]. Si bien el volumen es y será, un desafío significativo del Big Data se debe prestar mucha atención en todas las dimensiones del problema: Volumen, Variedad y Velocidad (conocidos como las 3Vs) [6].

El concepto de *streaming* en Big Data tiene algunas características notables, en estos sistemas los datos se reciben como una secuencia continua, infinita, rápida, en ráfagas, impredecible y que varía

en el tiempo [8]. El monitoreo (por ejemplo: tráfico de red, redes de sensores, cuidado de la salud, etc.), seguimientos de *clicks* en la web, transacciones financieras, detección de fraudes e intrusiones son algunas aplicaciones de *streaming* de Big Data [5]. Todos estos productores de datos que generan el *streaming* a menudo se encuentran distribuidos y con capacidades de procesamiento y memoria limitados.

En tareas de extracción de conocimiento dos pasos muy importantes son la selección de *features* (o atributos) y las tareas de mining de datos [12]. Estos problemas en entornos de *streaming* de datos siguen siendo un gran desafío. Una característica importante en selección de *features* es la habilidad para manejar grandes volúmenes de datos [14, 15]. Gran parte de las publicaciones existentes en la Web arriban como *streaming* (documentos, imágenes, contenidos multimedia, etc.), detectar un subconjunto de *features* útiles en estos flujos de datos en una tarea compleja debido a limitaciones de memoria, tiempos de respuesta, etc. [9, 18, 17].

Además del problema de selección de *features*, hay cuestiones enmarcadas dentro de las tareas de *streaming* mining para extracción de conocimiento. La problemática aquí radica en que los patrones de datos evolucionan continuamente y se torna necesario diseñar algoritmos de minería para tener en cuenta los cambios en la estructura subyacente del *streaming* de datos [3, 2, 16]. Incluso la distribución subyacente puede cambiar en el tiempo lo que genera que algunos modelos ya no sigan siendo válidos. Esto hace que las soluciones de los problemas sean aún más difíciles desde un punto de vista algorítmico y computacional [7].

En este trabajo se proponen diversas líneas de investigación sobre los temas mencionados, con aplicaciones en flujos de datos y problemas reales. Se abordan tanto problemas de mejoras de rendimiento ante distintos niveles de exigencia de precisión como también la escalabilidad de los diferentes abordajes a datos reales.

Líneas de I+D

En este proyecto se inician líneas de I+D relacionadas principalmente con el análisis y adaptación de algoritmos de aprendizaje automático en entornos de *Streaming* de datos. Puntualmente se está trabajando con árboles de decisión adaptativos en aplicaciones de cache de consultas sobre motores de búsqueda. Por otro lado, se está trabajando en profundizar los conocimientos en selección de atributos sobre *streaming* de datos. Por último, se están analizando diferentes opciones de topologías con Storm Apache¹ para la resolución de problemas de ETL. A continuación se hace una descripción somera de las líneas de I+D.

a. Árboles de decisión adaptativos

Los árboles de decisión adaptativos, o *Hoeffding Adaptive Tree* (HAT) [3], son una variante de *Hoeffding Tree* que utilizan ventanas deslizantes para mantener ajustado el árbol, sin embargo no requiere que el usuario le especifique el tamaño de ventana a utilizar. Esto se debe a que el tamaño de ventana óptimo se calcula individualmente para cada nodo, utilizando detectores de cambios y estimadores llamados ADWIN [4]. Los resultados preliminares de aplicar HAT en el dominio de gestión de caché de consultas en motores de búsqueda han sido muy alentadores.

Nuestro siguiente paso es realizar experimentos en otros dominios como, por ejemplo: análisis de sentimientos en redes sociales. De esta manera esperamos cuantificar la respuesta de éste método de extracción de conocimiento en contextos sometidos a constantes cambios de conceptos.

b. Selección de atributos

Enormes fuentes de datos se generan continuamente a partir de fuentes tales como redes sociales, difusión de noticias, etc., y típicamente estos datos se encuentran en espacios de alta dimensión (como el espacio de vocabulario de un idioma). La selección de atributos, o *Feature Selection* (FS), para

¹<http://storm.apache.org/>

descubrimiento de conocimiento, ha representado un gran desafío durante los últimos años tanto en estadística como en problemas de aprendizaje automático [17]. En el contexto de *streaming* de datos de gran volumen, detectar un subconjunto de *features* que sean relevantes es un problema muy difícil de resolver por las siguientes razones:

1. El *stream* de datos puede ser infinito, por lo que cualquier algoritmo que trabaje *off-line* que intente almacenar toda la secuencia para el análisis se quedará sin memoria.
2. La importancia de los atributos cambia dinámicamente con el tiempo debido a la volatilidad de un concepto, un atributo importante pueden volverse insignificantes y viceversa.
3. Para varias aplicaciones en línea, es importante obtener el subconjunto de características en tiempo casi real.

Esta línea de investigación es complementaria de la anterior y se propone analizar la precisión en métodos supervisados para selección de *features* de fuentes de datos no estructuradas provenientes de redes sociales. Este abordaje va a permitir trabajar la dinámica de los cambios de conceptos en *streaming* de datos en tareas de clasificación.

c. Topologías Storm

En esta línea de investigación se propone trabajar con herramientas de *Stream Processing Engine* (SPE) para probar los diferentes algoritmos de aprendizaje automático (como HAT) y las diferentes estrategias de selección de atributos. Un SPE es un framework que tiene por objetivo abordar el desafío de procesar grandes volúmenes de datos, en tiempo real y sin requerir el uso de código específico. Sobre los SPE es posible implementar algoritmos de *machine learning* para extraer conocimiento de los *streaming* de datos.

La idea de utilizar herramientas como Apache Storm es poder definir topologías de procesamiento de manera ágil para gestión de estadísticas necesarias para las etapas de selección de *features* y en

tareas de aprendizaje supervisado y no supervisado.

Resultados y Objetivos

El objetivo principal de la propuesta es estudiar, desarrollar, aplicar, validar y transferir modelos, algoritmos y técnicas que permitan construir herramientas y/o arquitecturas para abordar algunas de las problemáticas relacionadas con el tratamiento de información masiva utilizando algoritmos de aprendizaje automático de Big Data para dar respuestas en tiempo real. Se propone profundizar sobre el estado del arte y definir, analizar y evaluar nuevos enfoques sobre aprendizaje automático a partir de *streaming* de datos. En particular se estudiarán las siguientes líneas principales:

1. Estrategias de gestión *streaming* de datos masivos para determinar las mejores herramientas para extracción de features y resolución de los problemas clásicos de ETL en el contexto del real-time.
2. Evaluar la escalabilidad de los algoritmos tradicionales del área de aprendizaje automático a problemas de respuestas en tiempo real sobre *streaming* de datos masivos en diferentes dominios.
3. Elaboración de metodologías para el desarrollo de modelos en línea para toma de decisiones a partir de fuentes de información heterogénea.

Formación de Recursos Humanos

Este proyecto brinda un marco para que algunos docentes auxiliares y estudiantes lleven a cabo tareas de investigación y se desarrollen en el ámbito académico. Actualmente, se están dirigiendo dos trabajos finales correspondientes a la Lic. en Sistemas de Información (UNLu). Se espera dirigir al menos dos estudiantes más por año y presentar dos candidatos a becas de investigación.

Referencias

- [1] AGGARWAL, C. C., ASHISH, N., AND SHETH, A. The internet of things: A survey from the data-centric perspective. In *Managing and mining sensor data*. Springer, 2013, pp. 383–428.
- [2] BALDOMINOS, A., ALBACETE, E., SAEZ, Y., AND ISASI, P. A scalable machine learning online service for big data real-time analysis. In *Computational Intelligence in Big Data (CIBD), 2014 IEEE Symposium on* (2014), IEEE, pp. 1–8. 00017.
- [3] BIFET, A. Adaptive stream mining: Pattern learning and mining from evolving data streams. In *Proceedings of the 2010 conference on adaptive stream mining: Pattern learning and mining from evolving data streams* (2010), Ios Press, pp. 1–212.
- [4] BIFET, A., AND GAVALDA, R. Learning from time-changing data with adaptive windowing. In *Proceedings of the 2007 SIAM international conference on data mining* (2007), SIAM, pp. 443–448.
- [5] BIFET, A., AND MORALES, G. D. F. Big data stream learning with samoa. In *Data Mining Workshop (ICDMW), 2014 IEEE International Conference on* (2014), IEEE, pp. 1199–1202.
- [6] CHEN, M., MAO, S., AND LIU, Y. Big data: a survey. *Mobile Networks and Applications* 19, 2 (2014), 171–209. 00324.
- [7] GABER, M. M., ZASLAVSKY, A., AND KRISHNASWAMY, S. Mining data streams: a review. *ACM Sigmod Record* 34, 2 (2005), 18–26.
- [8] GAMA, J. *Knowledge discovery from data streams*. CRC Press, 2010.
- [9] HUANG, H., YOO, S., AND PRASAD, S. Un-supervised Feature Selection on Data Streams. 1031–1040.
- [10] KHAN, N., YAQOUB, I., HASHEM, I. A. T., INAYAT, Z., MAHMOUD ALI, W. K., ALAM, M., SHIRAZ, M., AND GANI, A. Big data: survey, technologies, opportunities, and challenges. *The Scientific World Journal* 2014 (2014). 00028.
- [11] MARZ, N., AND WARREN, J. *Big Data: Principles and best practices of scalable real-time data systems*. Manning Publications Co., 2015.
- [12] PRUENKARN, R., WONG, K., AND FUNG, C. A review of data mining techniques and applications. *Journal of Advanced Computational Intelligence and Intelligent Informatics* 21, 1 (2017), 31–48.
- [13] SAFAEI, A. A. Real-time processing of streaming big data. *Real-Time Systems* 53, 1 (2017), 1–44. 00004.
- [14] TANG, J., ALELYANI, S., AND LIU, H. Feature selection for classification: A review. *Data Classification: Algorithms and Applications* (2014), 37.
- [15] WANG, J., ZHAO, P., HOI, S. C., AND JIN, R. Online feature selection and its applications. *IEEE Transactions on Knowledge and Data Engineering* 26, 3 (2014), 698–710.
- [16] WANG, L. Machine learning in big data. *International Journal of Advances in Applied Sciences* 4, 4 (2016), 117–123. 00048.
- [17] WU, X., YU, K., WANG, H., AND DING, W. Online streaming feature selection. In *Proceedings of the 27th international conference on machine learning (ICML-10)* (2010), Citeseer, pp. 1159–1166.
- [18] ZHICHAO, Y., AND CHUNYONG YIN, L. F. A Feature Selection Algorithm of Dynamic Data-Stream Based on Hoeffding Inequality. 92–95.

Minería de Datos y Visualización de Información

**Esteban Schab, Ramiro Rivera, Luciano Bracco, Facundo Coto,
Patricia Cristaldo, Lautaro Ramos, Natalia Rapesta, Juan Pablo Núñez,
Soledad Retamar, Carlos Casanova, Anabella De Battista**

Grupo de Investigación en Bases de Datos, Departamento Ingeniería en Sistemas de Información,
Fac. Reg. Concepción del Uruguay, Universidad Tecnológica Nacional
Entre Ríos, Argentina

{schabe, riverar, braccol, cotof, cristaldop, ramosl, rapestan, nunezjp, retamars,
casanovac,debattistaa}@frcu.utn.edu.ar

Norma Edith Herrera

Departamento de Informática, Universidad Nacional de San Luis, San Luis, Argentina
nherrera@unsl.edu.ar

Resumen

El procesamiento y análisis de las grandes cantidades de datos que se producen en la actualidad, posibilitan el hallazgo de patrones y tendencias ocultos en los mismos, que impacta directamente en la toma de decisiones en diversas áreas de estudios.

Se generan datos a gran velocidad y en grandes cantidades que requieren ser procesados para poder actuar de manera rápida. Como es el caso de la observación de turnos que se generan en entidades bancarias, donde hay momentos del día en que se requiere modificar los esquemas de atención, según la afluencia de determinadas categorías de clientes o el incremento de demandas de determinados servicios.

Existen numerosas técnicas de minería de datos aplicables a distintos casos de análisis de datos, que permiten obtener ventajas de esas grandes cantidades de datos almacenados.

En este artículo se presentan los tópicos de interés del proyecto *Minería de Datos: su aplicación a repositorios de datos masivos*, en el que se investigan tanto temas de minería de datos, como de visualización de información, como herramienta para representar de manera eficiente los resultados obtenidos.

Palabras clave: Minería de datos, streaming de datos, gestión de proyectos, visualización, scrapping.

Contexto

El presente trabajo se desarrolla en el ámbito del proyecto *Minería de Datos: su aplicación a repositorios de datos masivos (UT13781TC)* del Grupo de Investigación en Bases de Datos, perteneciente al Departamento Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Concepción del Uruguay.

1. Introducción

El Descubrimiento de Conocimiento en Bases de Datos consiste en el análisis automático exploratorio y modelado de grandes repositorios de datos e involucra áreas de conocimiento como inteligencia artificial, aprendizaje automático, estadística, sistemas de gestión de base de datos, técnicas de visualización de datos y medios que apoyan toma de decisiones.

La Minería de Datos involucra e integra técnicas de diferentes disciplinas tales como tecnologías de bases de datos y data warehouse, estadística, aprendizaje de máquina, computación de alta performance, computación evolutiva, reconocimiento de patrones, redes neuronales, visualización de datos, recuperación de información, procesamiento de imágenes y señales, y análisis de datos espaciales o temporales.

Como un subárea específica de la Minería de Datos se puede citar al Data Stream Mining,

que es el proceso de extraer conocimiento en estructuras de datos continuas y con rápidas transiciones [1]. Los data streams son datos que se generan de forma continua y a altas velocidades. El origen de dichos datos puede provenir de diversas fuentes, como registros generados por clientes que utilizan aplicaciones móviles, transacciones electrónicas, logs de navegación de una red de datos, información de redes sociales, datos provenientes de dispositivos wearables, entre muchos otros ejemplos.

El procesamiento de estos datos debe realizarse de forma secuencial y gradual registro por registro, o bien en ventanas de tiempo graduales. Los resultados de dicho procesamiento se utilizan para una amplia variedad de tipos de análisis, como correlaciones, agregaciones, filtrado y muestreo. Las conclusiones obtenidas a partir de dicho análisis aporta a las empresas visibilidad de numerosos aspectos del negocio y de las actividades de sus clientes, como la tasa de uso de un servicio, la actividad de un servidor, la ubicación geográfica de un móvil, personas o mercadería, la afluencia de determinado tipo de clientes, entre otros aspectos, y les permite responder con rapidez ante cualquier situación que surja. Por ejemplo, un banco podría analizar el incremento de determinada categoría de clientes en un momento dado y responder rápidamente habilitando más puestos de atención al cliente.

2. Líneas de Investigación, Desarrollo e Innovación

La línea de trabajo principal de nuestro proyecto de investigación es el estudio de técnicas de Minería de Datos aplicables a datos estructurados y no estructurados, atendiendo principalmente a su eficiencia y escalabilidad [2]. Actualmente se trabaja en la aplicación de técnicas de minería de datos en aplicaciones específicas como el procesamiento de streams de datos, *el análisis bibliométrico* y también en la temática *Agenda setting*.

2.1. Análisis Bibliométrico

Se trabaja en análisis bibliométrico tradicional y alternativo, midiendo el impacto de publicaciones científicas en sus distintas modalidades de difusión. Actualmente se está elaborando un análisis cuantitativo de publicaciones de autores de instituciones argentinas en la bases de datos SCOPUS de Elsevier [3], accedida desde la Biblioteca Electrónica de Ciencia y Tecnología del Ministerio de Ciencia, Tecnología e Innovación Productiva de la Nación y a través de la API proporcionada por SCOPUS [4], utilizando scripts desarrollados en lenguaje R [5]. Para las búsquedas se establece una palabra o frase clave. En algunos casos se aplicó como filtro que las publicaciones correspondiesen a Argentina para identificar y reunir los trabajos en los que al menos uno de los autores incluyera la mención de una institución argentina en los datos de afiliación institucional, a fin de poder comparar con la cantidad de publicaciones del resto del mundo.

Se comenzó trabajando en índices tradicionales de análisis de publicaciones (conocido como Modo 1), que se basan principalmente en analizar las publicaciones realizadas en revistas con referato pagas revistas pagas con referato. En la actualidad, se está trabajando en la actualidad en el denominado Modo 2, estudiando instituciones de pertenencia de los artículos provenientes de Scopus y además de en fuentes como Altmetric, en este último caso se busca información publicada en blogs de ONGs, institucionales, etc. Está previsto realizar un análisis del impacto de publicaciones en redes sociales como Facebook o Twitter.

2.2. Agenda setting

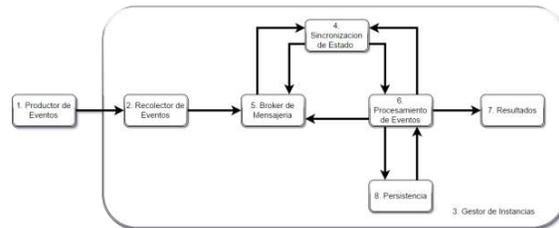
El término Agenda Setting hace referencia a la influencia que tienen los medios de comunicación en la fijación de temas en la opinión pública [6].

Se comenzó a realizar un trabajo de medición de los efectos de la instalación de asuntos en la agenda pública tomando como base artículos escritos sobre diferentes temáticas en medios digitales de relevancia para

determinar los tópicos que tratan y luego analizar su difusión en redes sociales empleando técnicas de minería de textos y procesamiento de lenguaje natural [7].

2.3. Sistema de procesamiento de streaming de datos

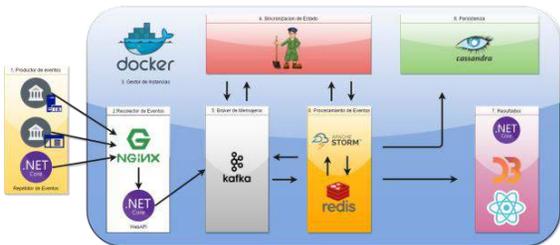
Otra parte de los esfuerzos del grupo se hallan abocados al estudio del procesamiento de datos en streaming, tema que cobra cada vez más protagonismo, tanto a nivel académico como por su capacidad de aportar a la Inteligencia de Negocios de las organizaciones. Ya no es suficiente con ser capaces de procesar grandes cantidades de datos extraídos de repositorios o generados por las organizaciones, sino que deben ser procesados de manera rápida, o en “real time”, además de generar información precisa. Los datos en streaming pueden provenir de diversas fuentes, como archivos de registros generados por los clientes que utilizan sus aplicaciones móviles o web, compras electrónicas, información de redes sociales, operaciones bursátiles o servicios geoespaciales. Algunos casos de aplicación del análisis de streaming de datos son la detección de fraudes, monitoreo de sistemas, intercambios comerciales y demás. El procesamiento de streams en tiempo real está diseñado para analizar y actuar en función de información a medida que la misma se genera, mediante el uso de consultas continuas (consultas del tipo SQL que operan sobre ventanas temporales e informacionales) [8]. Esto requiere un cambio de paradigma en cuanto al almacenamiento, obtención y procesamiento de la información. Las “bases de datos tradicionales” no fueron concebidas para este propósito por lo que debe hacerse uso de otras herramientas que otorguen la potencia y versatilidad que se requieren en este esquema de aplicaciones. Con este fin se trabajó en la definición de una arquitectura capaz de procesar streams de datos. Se propuso una arquitectura genérica, la cual representa los componentes necesarios para realizar la captura y procesamiento de los streams y sus interacciones.



Los componentes de la arquitectura deben ser capaces de interconectarse entre sí, proveer una alta tolerancia a fallas y permitir una escalabilidad elevada. En este esquema se pensó en la utilización de herramientas basadas en Software Libre, que se hallan respaldadas en el conocimiento colectivo de su comunidad. La arquitectura propuesta permitirá realizar el procesamiento de datos en streaming y será capaz de responder con una latencia máxima de 30 segundos a partir de un volumen de 100 eventos/seg.

A partir de la arquitectura genérica propuesta y los requerimientos antes mencionados, se planteó una Arquitectura para el Procesamiento de Streams de Datos que se encuentra en desarrollo haciendo uso de las siguientes herramientas:

- *Kafka*: broker de mensajería, utilizado para centralizar la recepción de información sobre los eventos que se produzcan.
- *Zookeeper*: mecanismo de sincronización distribuido, que mantiene el estado y configuración de las demás piezas de software del sistema.
- *Docker*: tecnología de contenerización, distribución de aplicaciones y virtualización, con el fin de garantizar sencillez de despliegue y posibilidad de escalado de la arquitectura.
- *Storm*: sistema distribuido de procesamiento de eventos en streaming, capaz de definir los “camino” y “transformaciones” que sufren los eventos para poder extraer datos de interés para la organización.
- *Redis*: Base de datos NoSQL, del tipo clave-valor, utilizada para permitir la reconfiguración del sistema sin necesidad de Down-times.



2.4. Transferencia tecnológica a industrias de la zona

En el marco de un convenio con una empresa local de desarrollo de software, se está realizando el desarrollo del prototipo para el procesamiento y posterior análisis de streaming de datos provenientes de las aplicaciones que esta empresa comercializa en bancos, con el objetivo de ofrecer información agregada para la toma de decisiones y que pueda retroalimentar dicha aplicación para automatizar ciertas decisiones a futuro.

En conjunto con el Grupo de Investigación y Desarrollo en Innovación y Competitividad del Departamento Licenciatura en Organización Industrial, de la FRCU-UTN, se encuentra en desarrollo el proyecto “Fortalecimiento de la Gestión productiva integral en PyMEs del sector metalmeccánico del Parque Industrial de Concepción del Uruguay, Entre Ríos”. Dicho proyecto ha resultado seleccionado en el marco de la convocatoria “Agregando Valor” (edición 2017) de la Secretaría de Políticas Universitarias de la Nación, orientada a la presentación de proyectos de vinculación tecnológica de alto impacto con la finalidad de transferir conocimientos y tecnologías innovadoras al sector socio-productivo nacional.

2.5. Visualización de datos

La generación y el almacenamiento de grandes volúmenes de información hacen que el mismo pase desapercibido y muchas veces se pierde la oportunidad de encontrar valor en ella. La visualización de datos es el proceso de representación de datos, en formato gráfico, de una manera clara y eficaz. Se convierte en una herramienta poderosa para el

análisis e interpretación de datos grandes y complejos, volviéndose un medio eficiente en la transmisión de conceptos en un formato universal [9, 10].

En este proyecto se trabaja en el análisis de técnicas y herramientas de visualización de datos, para mejorar los procesos de comunicación de resultados de las actividades que desarrolla el grupo. A partir de la generación de visualizaciones de dichos resultados, se logra una mejor comprensión de los datos. Entre las herramientas utilizadas actualmente se encuentran Tableau[11], Gephi[12], D3js[13], React D3[14] y Shiny[15].

2.6. Aplicación de metodología para la gestión de proyectos de Minería de Datos

En la gestión de las actividades de cada una de las líneas de investigación y desarrollo del proyecto se emplean fundamentos de metodologías ágiles. [16, 17] Partiendo de la propuesta metodológica de CRISP-DM [18] se realizó una adaptación empleando dichos fundamentos ágiles. Se espera poder formalizar dicha adaptación de CRISP-DM como una propuesta de metodología ágil para la gestión de proyectos de ciencia de datos.

3. Resultados obtenidos y esperados

Con este proyecto se espera lograr aplicaciones novedosas de técnicas y herramientas de minería de datos, en particular en áreas de estudio como bibliometría, la teoría de establecimiento de agenda. Además se espera obtener una herramienta eficiente en el análisis de datos en streaming. Todas estas iniciativas se desarrollan mediante la aplicación de una metodología ágil para proyectos de ciencias de datos.

4. Formación de Recursos Humanos

Este proyecto dio inicio a una nueva línea de investigación dentro del Grupo de investigación en Bases de Datos de la Fac. Reg. Concepción del Uruguay de la U.T.N..

Dos de los investigadores del proyecto están desarrollando tesis de maestría. En el proyecto colaboran dos becarios graduados con beca de iniciación a la investigación, que tienen previsto la realización de posgrados en el área temática del proyecto. Además participan en el proyecto cuatro becarios alumnos de la carrera Ingeniería en Sistemas de Información que inician su formación en la investigación, dos de ellos han realizado su Práctica Supervisada en el marco del proyecto.

5. Referencias

- [1] Khan, Latifur, and Wei Fan. 2012. "Tutorial: Data Stream Mining and Its Applications." In *Database Systems for Advanced Applications*, eds. Sang-goo Lee et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 328–329.
- [2] Larose Daniel T. *Discovering Knowledge in Data: An Introduction to Data Mining*. Wiley-Interscience, 2004.
- [3] SCOPUS. <http://www.scopus.com> Accedido 03/2018.
- [4] Scopus API. <https://goo.gl/mqpFpA> Accedido 03/2017.
- [5] R Project. <https://www.r-project.org/> Accedido 03/2018.
- [6] M. McCombs and D. Shaw. The agenda-setting function of mass media. *Public opinion quarterly*, 36(2):176–187, 1972.
- [7] Yeoul Kim, Suin Kim, Alejandro Jaimes, and Alice Oh. A computational analysis of agenda setting. In *Proceedings of the 23rd International Conference on World Wide Web (WWW '14 Companion)*, 323-324, 2014.
- [8] Tyler Akidau. The world beyond batch: Streaming 101. <https://goo.gl/xhPVZQ>. Accedido 03/2018.
- [9] Sadiku, Matthew (2016). Data Visualization. *International Journal of Engineering Research And Advanced Technology(IJERAT)*.Volume. 02. Issue.12.
- p. 11-16.
- [10] Finch, Jannette L., and Angela R. Flenner. 2017. "Using Data Visualization to Examine an Academic Library Collection." *College & Research Libraries* 77(6). <https://goo.gl/fAeW3w> (March 18, 2018).
- [11] Tableau. <https://www.tableau.com/es-es> Accedido 03/2018.
- [12] Gephi. <https://gephi.org/> Accedido 03/2018.
- [13] Data-Driven Documents. <https://d3js.org/> Accedido 03/2018.
- [14] React D3. <http://www.reactd3.org/> Accedido 03/2018.
- [15] Shiny from R Studio. <https://shiny.rstudio.com/> Accedido 03/2018.
- [16] Ken Schwaber and Jeff Sutherland. *The scrum guide*. Scrum Alliance, 2011, vol. 21.
- [17] Manifesto for Agile Software Development. Agile Alliance. <https://goo.gl/xRFCVL> . Accedido 03/2018.
- [18] Chapman, Clinton, Kerber, Khabaza, Reinartz, Shearer, & Wirth. *CRISP-DM 1.0 Step-by-step data mining guide*. 2000.

POLIMORFISMOS DE NUCLEÓTIDOS SIMPLES RELACIONADOS AL RIESGO DE ENFERMEDADES: CLASIFICACIÓN AUTOMÁTICA DE ESTUDIOS EPIDEMIOLÓGICOS DE TIPO CASO-CONTROL UTILIZANDO TÉCNICAS DE MINERÍA DE TEXTO

Mónica R. Mounier^{1,2,a}, Karina B. Acosta^{1,b}, Fabián Favret^{2,c}, Eduardo Zamudio^{1,d}, Diego A. Godoy^{2,e}, Juan de Dios Benítez^{2,f}

1 - Facultad de Ciencias Exactas Químicas y Naturales, Universidad Nacional de Misiones, Félix de Azara 1552, Posadas, Misiones-Argentina. Tel: +54 (0376) – 4422186

2 - Universidad Gastón Dachary, Avda. López y Planes 6519, Posadas, Misiones-Argentina. Tel: +54 (0376) -4438677

^amonicamounier@fceqyn.unam.edu.ar, ^bacostakb2505@gmail.com, ^cfabianfavret@citic.ugd.edu.ar, ^deduardo.zamudio@fceqyn.unam.edu.ar, ^ediegodoy@citic.ugd.edu.ar, ^fjuan.benitez@citic.ugd.edu.ar

RESUMEN

En este trabajo de investigación se presenta una herramienta bioinformática que permite clasificar automáticamente artículos científicos referentes a estudios epidemiológicos de tipo caso-control concernientes a Polimorfismos de Nucleótidos Simples (SNPs), presentes en genes, y su asociación a distintos tipos de cáncer, y otras enfermedades genéticas de interés para el experto mediante la utilización de técnicas de minería de texto (MT), así también como la implementación del meta-estimador *Bagging* para tres técnicas de clasificación: *Support Vector Machine* (SVM), *K-Nearest Neighbors* (KNN), y *Naives Bayes* (NB). La clasificación se realiza a partir de los metadatos de dichos artículos, los cuales están disponibles en el *National Center for Biotechnology Information* (NCBI).

Palabras clave: *Bioinformática, Minería de Textos, Meta-estimadores, Polimorfismos, Clasificación Automática, Estudios Epidemiológicos.*

CONTEXTO

El trabajo presentado en este artículo tiene como contexto marco el proyecto de investigación denominado “Simulación en las Tics: Diseño de Simuladores de Procesos de Desarrollo de Software Ágiles y Redes De Sensores Inalámbricos para la Industria y la Academia”, registrado actualmente en la Secretaría de Investigación y Desarrollo de la Universidad Gastón Dachary (UGD) con el número Código IP A07003 y radicado en el Centro de Investigación en Tecnologías de la Información y Comunicaciones de dicha universidad.

El mismo fue incorporado como proyecto aprobado en el llamado a presentación interna de la UGD de proyectos de investigación N°7 mediante la Resolución Rectoral 07/A/17 y es una continuidad del Proyecto Simulación como herramienta para la mejora de los procesos de software desarrollados con metodologías ágiles utilizando dinámica de sistemas, R.R. UGD N° 18/A/14 y R.R. UGD N° 24/A/15.

1. INTRODUCCIÓN

En la última década se ha visto un enorme crecimiento en la cantidad de datos biomédicos experimentales y computacionales, específicamente en las áreas de genómica y proteómica. Este crecimiento ha aumentado el número de publicaciones biomédicas referentes a discusiones sobre hallazgos. Debido a ello, hay un gran interés por parte de la comunidad científica en herramientas de minería para ayudar a clasificar la abundante documentación disponible, a fin de encontrar datos relevantes y útiles para tareas de análisis específicas [1].

Particularmente, los investigadores en el área biomédica, en adelante llamados expertos, presentan, como resultado de sus investigaciones y hallazgos, artículos científicos no estructurados. Dichos artículos son utilizados posteriormente por otros expertos para el estudio, diagnóstico, tratamiento y/o prevención de enfermedades. El inmenso cuerpo y rápido crecimiento del corpus biomédico han llevado a la aparición de un gran número de técnicas de MT destinadas a la extracción automática de conocimiento [2].

Actualmente, existe la necesidad de disponer de una herramienta que permita clasificar automáticamente artículos científicos referentes a estudios epidemiológicos de tipo caso-control, que reflejen la asociación de SNPs, presentes en genes, y su asociación a enfermedades específicas. La fuente principal de metadatos de los artículos a clasificar se encuentra disponible en el NCBI [3]. Hoy en día, dicha clasificación es realizada manualmente por el experto, lo cual resulta notablemente ineficiente, dada la cantidad de tiempo necesaria para llevarla a cabo, sumado a las constantes actualizaciones de la bibliografía disponible.

Los SNPs son variaciones de la secuencia de ácido desoxirribonucleico (ADN) que se producen cuando se altera un solo nucleótido (A, T, C o G) en el genoma humano. Los SNPs representan alrededor del 90% de toda la variación genética humana, que se producen cada 100-300 bases a lo largo del genoma humano (3 mil millones de bases), aunque su densidad varía entre las regiones [4].

La bioinformática consiste en la investigación, desarrollo y/o aplicación de herramientas computacionales y enfoques para ampliar el uso de datos biológicos, médicos, de comportamiento o de salud, incluidas las de adquirir, almacenar, organizar, archivar, analizar y visualizar esos datos [5].

La MT procesa la información no estructurada y extrae índices numéricos desde el texto, a partir de lo cual hace a la información accesible para varios algoritmos de minería de datos. Básicamente, la MT convierte el texto en números los cuales pueden luego ser incluidos en otros análisis [6].

Las tecnologías de MT básicas pueden ser utilizadas individualmente o en forma conjunta, dependiendo del problema de minería a resolver. Dichas tecnologías son las siguientes [7]:

- Recuperación de información;
- Extracción de información;
- Categorización o clasificación;
- *Clustering*;
- Generación Automática de Resúmenes.

En la recuperación de información, los documentos son generalmente identificados por un conjunto de términos o palabras clave que son usadas colectivamente para representar su contenido [8]. Existen muchas técnicas de representación de textos, siendo algunas de las más utilizadas las siguientes [9, 10]:

- *Inverse Document Frequency* (IDF);
- *Term Frequency – Inverse Document Frequency* (TF-IDF).

El método *Bagging* es un meta-estimador en el cual los clasificadores individuales son entrenados en paralelo a partir de muestras con reemplazo obtenidas aleatoriamente del mismo conjunto de entrenamiento. Para construir un comité de n clasificadores (siendo n el número de clasificadores a utilizar), se debe elegir la forma de combinar los resultados de dichos clasificadores, siendo el método más simple el voto mayoritario en el cual la categoría asignada a un documento es aquella que fuera elegida por la mayoría de los clasificadores, donde n debe ser un número impar [11]. En la Figura 1 se muestra el esquema de funcionamiento del meta-clasificador *Bagging*.

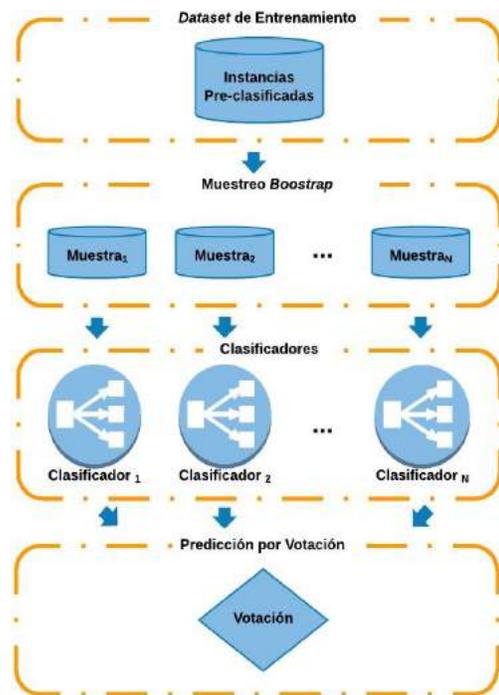


Figura 1. Esquema del Meta-estimador Bagging

2. LÍNEAS DE INVESTIGACIÓN

En esta investigación el objetivo es el desarrollo e implementación de una herramienta bioinformática de clasificación automática de estudios epidemiológicos de tipo caso-control referentes a SNPs relacionados a distintos tipos de cáncer, y otras enfermedades genéticas de interés para el experto utilizando técnicas de MT, a partir de sus metadatos.

Objetivos Específicos

- Identificar los metadatos de recuperación y clasificación necesarios para la elaboración de la herramienta bioinformática;
- Seleccionar las técnicas adecuadas para el tratamiento de los datos recuperados;
- Obtener una base de instancias pre-clasificadas por el experto a partir de la información relevante recuperada a través del NCBI relacionada a estudios epidemiológicos de tipo caso-control referentes SNPs
- Diseñar e implementar la herramienta bioinformática para la clasificación automática de estudios epidemiológicos de tipo caso-control referentes SNPs relacionados a enfermedades de interés para el experto

3. RESULTADOS

Para la elaboración del *dataset* fueron consideradas aleatoriamente los metadatos de 198 citas bibliográficas de artículos científicos de interés para el experto, clasificadas por el mismo en dos categorías: “Asociados” (169 artículos) y “No Asociados” (29 artículos). Siendo un problema intrínseco el desbalanceo de las

clases, dado que la mayoría de los estudios reflejan asociaciones de los SNP a las enfermedades y no lo contrario.

Para el presente trabajo ha sido realizada la adaptación de una metodología CRISP_DM [12], así también como la metodología propuesta en un trabajo de investigación aplicado a MT [13]. Las etapas que componen la metodología son:

1. Recuperación de metadatos;
2. Pre-procesamiento de metadatos;
3. Representación de los datos;
4. Descubrimiento de conocimiento.

En la Figura 2 se presenta la herramienta que ha sido desarrollada en el presente trabajo. La misma está conformada por los siguientes módulos:

1. Módulo de consulta;
2. Módulo de recuperación;
3. Módulo de pre-procesamiento;
4. Módulo de clasificación;
5. Módulo de visualización;
6. Módulo de retroalimentación.

Las herramientas bioinformáticas utilizadas fueron:

- Biopython: Herramientas de libre acceso para la biología computacional y bioinformática en Python.
- E-utilities: Conjunto de programas que proporcionan una interfaz de consulta de Entrez del NCBI.
- genenames.org Rest Web Service: Servicio web proporcionado por el HUGO Gene Nomenclature Committee.

Para la representación de los metadatos de los artículos fue utilizado TF-IDF de los unigramas de los mismos.

Para la clasificación fue utilizado el meta-estimador *Bagging*, por ser el más adecuado para *dataset* desbalanceados [14], para tres técnicas de clasificación: *Support Vector Machine* (SVM), *K-Nearest Neighbors* (KNN) y *Naives Bayes* (NB), utilizando el 60 % del *dataset* para entrenamiento y el 40 % restante para validación.

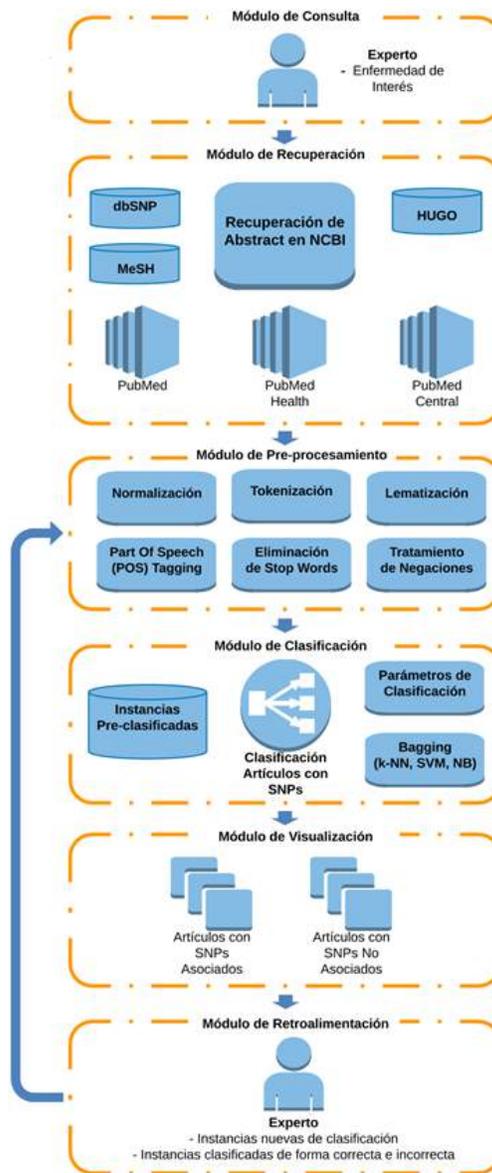


Figura 2. Esquema del Meta-estimador

En la Tabla 1 se presentan los resultados obtenidos de la clasificación para los tres meta-clasificadores utilizando las tres técnicas de clasificación planteadas, las cuales fueron entrenadas y validadas sobre el mismo subconjunto de datos, a fin de comparar el comportamiento de cada meta-clasificador.

Tabla 1. Resultados obtenidos

	<i>Bagging</i> SVM		<i>Bagging</i> KNN		<i>Bagging</i> NB	
Matriz de Confusión						
	A	NA	A	NA	A	NA
A	65	0	64	1	64	1
NA	10	4	13	1	8	6
Medidas de Evaluación						
Precisión	0.89		0.77		0.88	
Cobertura	0.87		0.82		0.89	
<i>F1-Score</i>	0.84		0.76		0.87	
Exactitud	0.97		0.94		0.98	

Donde A: Asociados; NA: No Asociados

Como puede verse en Tabla 1, los resultados obtenidos fueron superiores para el *Bagging* con NB, alcanzando una exactitud del 0.98 %.

4. FORMACIÓN DE RECURSOS HUMANOS

El trabajo corresponde a un trabajo de fin de carrera de grado de Ingeniería en Informática. El equipo de investigación está formado por un Doctor en Ciencias de la Computación, una Doctora en Ciencias Biológicas, ambos, docentes-investigadores de la UNaM, un Doctor en Tecnologías de la Información (TI) y Comunicación, un maestrando en TI, un maestrando en Redes de datos y ocho estudiantes en período de realización de trabajos finales de grado en el contexto de las carreras de Licenciatura en Sistemas de Información y de Ingeniería en Informática de la UGD. Actualmente, el número de tesinas de grado aprobadas en el contexto de este proyecto, es de cinco, y otras tres en proceso de desarrollo.

5. REFERENCIAS

[1] H. Shatkay and R. Feldman, "Mining the biomedical literature in the genomic era: an overview", *Journal of computational biology*, vol. 10, no. 6, pp. 821-855, 2003.

[2] F. Zhu et al., "Biomedical text mining and its applications in cancer research", *Journal of biomedical informatics*, vol. 46, no. 2, pp. 200-211, 2013.

[3] (2017, Feb.) National Center for Biotechnology Information (NCBI). [Online]. <http://www.ncbi.nlm.nih.gov>

[4] J. E. Lee, J. H. Choi, J. H. Lee, and M. G. Lee, "Gene SNPs and mutations in clinical genetic testing: haplotype-based testing and analysis", *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, vol. 573, no. 1, pp. 195-204, 2005.

[4] M. Huerta, G. Downing, F. Haseltine, B. Seto, and Y. Liu, "NIH working definition of bioinformatics and computational biology", US National Institute of Health, 2000.

[6] R. Agrawal and M. Batra, "A detailed study on text mining techniques", *International Journal of Soft Computing and Engineering*, vol. 2, no. 6, pp. 118-121, 2013.

[7] S. Dang and P. H. Ahmad, "Text Mining: Techniques and its Application", *International Journal of Engineering & Technology Innovations*, pp. 2348-0866, 2014.

[8] W. Zhang, T. Yoshida, and X. Tang, "A comparative study of TF* IDF, LSI and multi-words for text classification", *Expert Systems with Applications*, vol. 38, no. 3, pp. 2758-2765, 2011.

[9] K. S. Jones, "A statistical interpretation of term specificity and its application in retrieval", *Journal of documentation*, vol. 28, no. 1, pp. 11-21, 1972.

[10] K. S. Jones, "IDF term weighting and IR research lessons" *Journal of documentation*, vol. 60, no. 6, pp. 521-523, 2004.

[11] L. Breiman, "Bagging predictors" *Machine learning*, vol. 24, no. 2, pp. 123-140, 1996.

[12] P. Chapman et al., "CRISP-DM 1.0 Step-by-step data mining guide", 2000.

[13] C. Gálvez, "Minería de textos: la nueva generación de análisis de literatura científica en biología molecular y genómica" *Encontros Bibli: rev. eletrônica de bib. e ciência da inf.*, vol. 13, no. 25, pp. 1518-2924, 2008.

[14] G. Collell, D. Prelec, K. Patil, "A simple plug-in bagging ensemble based on threshold-moving for classifying binary and multiclass imbalanced data", *Neurocomputing*, vol. 275, pp. 330-340, 2018.

Métodos de Acceso para Bases de Datos Métricas

Jorge Arroyuelo, Maria E. Di Genaro, Damián Escudero, Alejandro Grosso, Verónica Ludueña,
Cintia Martínez, Nora Reyes

Dpto. de Informática, Fac. de Cs. Físico-Matemáticas y Naturales, Universidad Nacional de San Luis
{bjarroju, digeme, agrosso, vlud, nreyes}@unsl.edu.ar, escudero.damian.ez@gmail.com, cintiavmartinez@hotmail.com

Edgar Chávez

Centro de Investigación Científica y de Educación Superior de Ensenada, México

elchavez@cicese.mx

Karina Figueroa

Fac. de Cs. Físico-Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo, México

karina@fisimat.umich.mx

Rodrigo Paredes

Dpto. de Cs. de la Computación, Fac. de Ingeniería, Universidad de Talca, Chile

rapared@utalca.cl

Resumen

En la actualidad se ha incluido, en mayor o menor medida y en casi todo ámbito, a la ciencia de la computación. Esto a provocado todo tipo de requerimientos de usuarios de distinta índole, y desde campos muy disímiles. Para satisfacer estas demandas se deben desarrollar aplicaciones capaces de manipular eficientemente datos no convencionales muy dispares como: audio, huellas digitales, texto, video, imágenes, secuencias de ADN, etc. Además es necesario utilizar depósitos especializados y búsquedas no exactas sobre estos tipos de datos, porque las soluciones tradicionales no suelen enfrentar tales requerimientos.

Por otro lado, la gran cantidad de datos que se deben manipular para lograr respuestas adecuadas y eficientes, hace necesario un uso eficaz del espacio disponible, lo que implica que las estructuras utilizadas para acceder a este tipo de base de datos, deben ser *estructuras de datos conscientes de la jerarquía de memoria*. Un modelo en el cual se puede utilizar estructuras de datos especializadas o métodos de acceso que contemplen estos aspectos son las *Bases de Datos Métricas*. Con todas estas consideraciones en mente, esta investigación pretende contribuir a consolidar este nuevo modelo de bases de datos desde varias perspectivas.

Palabras Claves: bases de datos no convencionales, índices, lenguajes de consulta.

Contexto

La investigación que se realiza en este ámbito, está enfocada en lograr que las bases de datos destinadas a manipular datos no estructurados, no convencionales, alcancen la madurez de las bases de datos tradicionales. Esto incluye además, plantear nuevas arquitecturas del procesador que mejoren a

muy bajo nivel los administradores de estas bases de datos. Se busca así contribuir a distintos campos de aplicación: sistemas de información geográfica, robótica, visión artificial, diseño asistido por computadora, computación móvil, entre otros.

Los estudios que dieron lugar al presente trabajo se realizan en el marco del Proyecto *Tecnologías Avanzadas de Bases de Datos*, en la línea *Bases de Datos no Convencionales* de la Universidad Nacional de San Luis, en colaboración con investigadores de otros grupos de: Universidad Michoacana de San Nicolás de Hidalgo (México) y Centro de Investigación Científica y de Educación Superior de Ensenada (México). Este proyecto finalizó a fines de 2017 y su continuación está en proceso de evaluación.

Introducción

El advenimiento de la computación a todos los ámbitos de la sociedad, tanto el laboral, como el productivo, recreativo, científico, artístico, de la salud, etc., ha exigido el desarrollo de aplicaciones capaces de adaptarse tanto a estos nuevos entornos como a los diversos usuarios de las mismas. Para ello, las bases de datos han debido evolucionar hasta ser capaces de administrar todo tipo de datos y responder consultas sobre los mismos de una manera totalmente diferente a la tradicional, muchas veces más intuitiva. Estos avances se han visto reflejados en áreas como: comparación de huellas digitales, reconocimiento de voz, reconocimiento facial, bases de datos médicas, reconocimiento de imágenes, minería de

datos, recuperación de texto, biología computacional, clasificación y aprendizaje automático, etc.

El modelo de *espacios métricos* resulta adecuado para englobar ciertas características que comparten todas estas aplicaciones, a pesar de ser tan diversas. Formalmente un espacio métrico consiste de un universo de objetos y una función de distancia, definida entre ellos, que mide cuán diferentes son los objetos. Este escenario es propicio para resolver demandas tales como, ingresar una imagen a un buscador y esperar que éste muestre imágenes parecidas a la provista, o un trozo de canción y se encuentren las canciones similares a dicho trozo. Para este tipo de problemas, las búsquedas exactas carecen de sentido y son más naturales sobre estos tipos de datos las *búsquedas por similitud* que provee este modelo.

Para evitar la examinación secuencial de los datos al responder eficientemente a este tipo de búsquedas, se utilizan los *Métodos de Acceso Métricos* (MAMs). Sin embargo, es esencial su optimización, ya que la mayoría de estos métodos no admiten actualizaciones, ni están diseñados para soportar conjuntos masivos de datos y tampoco para resolver operaciones de búsquedas complejas.

El estudio de los lenguajes de consulta, es otra de las áreas exploradas, se busca incrementar su expresividad para formular consultas más precisas. Además, se busca caracterizar nuevas arquitecturas que permitan reducir el flujo de bits entre el procesador y la memoria en relación a la cantidad de datos utilizados por cada programa, para mejorar el desempeño en administradores de bases de datos (DBMS) a bajo nivel.

Líneas de Investigación y Desarrollo

Bases de Datos no Convencionales

Como se mencionó, se utilizarán los espacios métricos para modelizar aquellas bases de datos no convencionales, que administran videos, imágenes, texto libre, secuencias de ADN, audio, etc. En este ámbito es necesario responder eficientemente consultas por similitud, haciendo uso de MAMs y debido a lo costoso que resultan los cálculos de distancia, el número de cálculos realizados al crear el índice o al realizar búsquedas es usado como medida general de complejidad. Por ello, se analizan aquellos MAMs que han mostrado buen desempeño en las búsquedas, para optimizarlos, siendo conscientes para ello de la jerarquía de memorias.

En general, un espacio métrico consta de un universo \mathbb{U} y una función de distancia d y dada una base

de datos $X \subseteq \mathbb{U}$ y una consulta $q \in \mathbb{U}$, las consultas por similitud son de dos tipos: por *rango* o de *k-vecinos más cercanos* (k -NN).

Grafo de los k Vecinos

Una de las búsquedas por similitud tradicionales, la de los k vecinos más cercanos, es utilizada por aplicaciones como la predicción de funciones, la cuantificación y compresión de imágenes, la clasificación y aprendizaje automático, entre otras. En este caso, dado un elemento $u \in \mathbb{U}$ y sea $X \subseteq \mathbb{U}$ la base de datos, se recuperan los k elementos en $X - \{u\}$ que tengan la menor distancia d a u . Una generalización de esta búsqueda es la obtención de los k -vecinos más cercanos de *todos* los elementos de la base de datos (*All-k-NN*). La solución burda a este problema, comparar cada elemento de la base de datos con todos los demás, tiene una complejidad de n^2 cálculos de distancia, con $|X|=n$. Por ejemplo, una solución más eficiente resulta al preprocesar los datos por medio de un índice para reducir el número de cálculos de distancia en las búsquedas.

Entre las soluciones propuestas para espacios métricos generales, algunas se basan en la construcción del *Grafo de los k-vecinos más cercanos* (k NNG) [10], cuyo desempeño supera algunas de las técnicas clásicas. El k NNG indexa un espacio métrico y luego se emplea en la resolución de las consultas por similitud. Sin embargo, cuando la función de distancia es demasiado costosa de calcular, o si se tiene una base de datos masiva, el costo de la construcción de un índice, para luego obtener los vecinos más cercanos, puede resultar excesivo. Al igual que resolver consultas en espacios métricos de alta dimensión, donde muchas veces se requiere revisar casi todo el conjunto de datos sin importar la estrategia utilizada. Tanto para hacer frente a éstas situaciones, como para satisfacer los requerimientos de algunas aplicaciones que priorizan la velocidad sobre la precisión [11, 6, 12, 7], es que se consideran las *búsquedas por similitud aproximadas*. Es decir, se aceptan algunos “errores” en la respuesta, si con esto se mejora la complejidad de la misma. Teniendo en cuenta estas consideraciones, se han desarrollado algunas propuestas que construyen una aproximación del k NNG, que se denominó *Grafo de vecinos cercanos* (kn NG) [4], el cual conecta cada objeto de la base de datos con k vecinos *cercanos*, relajando la condición que los k vecinos devueltos sean los más cercanos a u de toda la base de datos. Entonces, se puede perder alguno muy cercano y en su lugar devolver otro un poco más lejano.

En una primera aproximación, se consideró un caso particular del k nNG, cuando $k = 1$, y se obtuvo el 1nNG, el grafo que conecta a cada elemento con un elemento cercano de la base de datos, que puede ser, o no, su vecino más cercano. Entonces, aprovechando el profundo conocimiento que se tiene de *DiSAT*, se propuso en [4] un enfoque novedoso al problema, que utiliza la información obtenida durante la construcción del índice para construir el 1nNG; cada objeto es vinculado con el elemento más cercano de la base de datos con el que se comparó durante la construcción del *DiSAT*. Esto retorna una aproximación del 1nNG, la cual, pese a que es bastante buena, puede mejorarse mediante reconstrucciones adicionales. Esta propuesta permite recuperar el 1nNG con bajo costo, una muy buena precisión y un error bajo, logrando un buen compromiso calidad/tiempo, y llamativamente *sin realizar ninguna búsqueda*.

Otras propuestas abordadas resuelven el k nNG sin recurrir a ningún índice, ni siquiera a su construcción. Se plantean distintas maneras de seleccionar muestras de la base de datos, a las que se le calculan sus vecinos más cercanos, y diferentes formas de utilizar la información conseguida en ese proceso, para calcular vecinos aproximados para el resto de los objetos; utilizando propiedades de la función de distancia, como la desigualdad triangular. Estos planteos resultan muy prometedores.

Métodos de Acceso Métricos

Como se dijo anteriormente, una de las optimizaciones necesarias a los MAM's es el dinamismo. Por ejemplo, considerando el *Árbol de Aproximación Espacial (SAT)*, un índice con muy buen desempeño en espacios de mediana a alta dimensión, pero totalmente estático, se desarrolló el *Árbol de Aproximación Espacial Dinámico (DSAT)* [9] que permite realizar inserciones y eliminaciones, conservando muy buen desempeño en las búsquedas, pero que agrega un parámetro a sintonizar. El *Árbol de Aproximación Espacial Distal (DiSAT)* [5], una variante también estática del *SAT* y sin parámetros, logra optimizar las búsquedas respecto de ambos (*SAT* y *DSAT*). Por ello, se ha propuesto la *Foresta de Aproximación Espacial Distal (DiSAF)* [3], que es dinámica, para memoria principal y que para lograr mejorar al máximo su desempeño, aplica la técnica de dinamización de Bentley y Saxe al *DiSAT* y aprovecha el profundo conocimiento que se tiene sobre la aproximación espacial.

Sin embargo, muchas veces los índices no caben en memoria principal, ya sea porque adminis-

tran una base de datos masiva, o porque los objetos de la misma son muy grandes. Entonces surge la necesidad de diseñar índices para memoria secundaria. Muchos de estos índices se basan en “agrupar elementos”; y para analizar cuán buenos son los agrupamientos que logran, se pueden utilizar estrategias de optimización basadas en heurísticas bioinspiradas. Teniendo esto en consideración, se han diseñado dos nuevos índices basados en la *Lista de Clusters(LC)* [6] que son totalmente dinámicos, es decir, admiten inserciones y eliminaciones de elementos y están especialmente diseñados para trabajar sobre grandes volúmenes de datos [9]. La *Lista de Clusters Dinámica (DLC)*, tiene buen desempeño en espacios de alta dimensión, con buena ocupación de página y operaciones eficientes tanto en cálculos de distancia como en operaciones de I/O. Sin embargo, las búsquedas en ella deben recorrer completamente la lista de centros de los clusters, elevando los costos. El *Conjunto Dinámico de Clusters (DSC)*, también mantiene los clusters en memoria secundaria, pero organiza los centros de clusters en un *DSAT* en memoria principal, permitiendo que las búsquedas realicen menos cálculos de distancia y accedan a menos páginas/clusters. La información de ese *DSAT* también se aprovecha en las inserciones, mejorando los costos de las operaciones en cálculos de distancia y manteniendo los bajos costos de acceso a disco. Ambos, *DLC* y *DSC*, han demostrado tener una razonable utilización de páginas de disco y son competitivas respecto a las alternativas representativas del estado del arte.

Otro aspecto a considerar en este caso es la calidad de los clusters generados. Por lo tanto, una variante que se está considerando para la *DSC* es que en lugar de insertar los elementos en el índice a medida que van llegando, se puede demorar la incorporación de cada nuevo elemento a un cluster hasta tener varios elementos y poder determinar así un mejor agrupamiento de los elementos. Esto permite además reducir el costo de construcción del índice, porque se realiza una escritura de un cluster en disco luego de varias inserciones y además implícitamente puede mejorar los costos de búsqueda al lograr clusters más compactos y que aseguran una total ocupación de la página del disco, achicando así el tamaño del archivo y reduciendo así los tiempos de acceso.

Algunas aplicaciones requieren que las respuestas sean aún más rápidas, aunque sea a costa de perder algunos elementos: se intercambia precisión (desvolviendo sólo algunos objetos relevantes) por ve-

locidad en la respuesta. Este tipo de búsquedas se denominan *aproximadas*. Para conjuntos de datos masivos, las búsquedas por similitud aproximadas permiten obtener un buen balance entre el costo de las búsquedas y la calidad de la respuesta obtenida. El *Algoritmo Basado en Permutaciones (PBA)* [2], es uno de los mejores representantes de este tipo de consultas, logrando una respuesta de alta calidad a un bajo costo. Por ello, se ha diseñado la *Lista Dinámica de Permutaciones Agrupadas (DLCP)* [8], que combina *LC* con *PBA*, es dinámica y para memoria secundaria. Este índice agrupa por distancia entre las permutaciones de los objetos, en lugar de por distancia entre objetos y se le puede indicar cuántos cálculos de distancia y/o operaciones de I/O utilizar, para obtener una respuesta rápida, aunque menos precisa. Además, se están considerando nuevas variantes para obtener mejores resultados.

Arquitecturas de Procesadores Orientadas a Bases de Datos

La arquitectura del procesador es la funcionalidad que se le provee al programador en lenguaje de máquina, modos de direccionamiento, operaciones, interrupciones y entrada-salida [1]. En ella se distinguen: la organización básica del flujo de datos y el control que se utilizan para alcanzar dicha funcionalidad (*implementación*) y la estructura física que se utiliza para materializar la implementación (*realización*). El lenguaje de máquina (LM) actual no es ni un lenguaje de aplicación ni un lenguaje de hardware, sino algo intermedio. Entonces, ¿por qué no interpretar directamente un lenguaje de alto nivel en lugar de compilar a un lenguaje intermedio? o ¿por qué no darle acceso directo al programador/compilador al hardware en lugar de restringirlo al LM? ¿En qué nivel debería estar el LM?

Se puede elegir la estrategia de “impulso hacia arriba”; es decir subir el nivel, para mejorar el desempeño de la máquina, además de facilitar el uso del lenguaje de máquina. Un aspecto a considerar en este caso es el tráfico de bits y la forma usual de reducirlo es tener una arquitectura que haga lo más posible con cada búsqueda de instrucción, abandonando la arquitectura de bajo nivel y yendo tan alto como el software lo permita. El otro aspecto es explotar la concurrencia, porque si una implementación conoce más sobre lo que debe ser hecho entonces es posible que a menudo realice varias acciones simultáneamente. El implementador posee varias técnicas para aumentar la concurrencia: paralelizar, segmentar (pipelining), adelantar, poner a un lado

(cache look-aside), adivinar y corregir (control and data prediction). La otra estrategia es considerar el “impulso hacia abajo”. Aún si todas las aplicaciones fueran escritas en lenguaje de alto nivel, hay razones para definir una arquitectura de computadora de nivel más bajo, pues existe conflicto de intereses entre usuario e implementador: el usuario desea expresar en forma simple y breve, haciendo uso del contexto, y el implementador desea que cada instrucción sea interpretada independientemente del resto.

Por lo tanto, es importante definir una arquitectura cuando se construye una computadora. En la actualidad la investigación sobre arquitecturas de procesadores ha sido desplazada por la de implementación de procesadores. La mayoría de los trabajos de investigación se dedican a mejorar técnicas de predicción (tanto de control como de datos), técnicas para sincronizar y comunicar procesadores (núcleos) mediante mensajes y/o memoria compartida. Muchas de estas técnicas de implementación surgieron en los años 60 y hoy se han incorporado a los diseños de microprocesadores actuales. Sin embargo, estas técnicas de implementación se podrían aplicar a todo tipo de arquitectura, desde una arquitectura RISC,¹ que intenta acercar el lenguaje de máquina al hardware del procesador, a una arquitectura que se aleje del hardware e intente disminuir el tráfico de bits entre procesador y memoria. El objetivo en esta área es plantear nuevas arquitecturas que minimicen el tráfico de bits entre el procesador y la memoria. Se está construyendo un simulador del set de instrucciones AMD-64 o x86-64, como “benchmark”, para evaluar el tráfico de bits, como Specint y Specfp para la arquitectura x86. Luego, se evaluará el tráfico de bits para la arquitectura propuesta sobre los mismos benchmarks, lo que implica construir tanto el simulador de la arquitectura como el compilador C para la misma. Finalmente, se pretende aprovechar el conocimiento adquirido para, desde bajo nivel, mejorar el desempeño de los DBMSs.

Otra línea de investigación en esta materia, se refiere al diseño de sumadores/restadores y multiplicadores de punto flotante, con vistas a generar una especificación en un lenguaje de descripción de hardware, que concuerde en parte con el estándar IEEE para números de simple precisión. La descripción debe ser sintetizable sobre lógica programable como por ejemplo los FPGA (Field Programmable Gate Array). Se considerarán distintas implementaciones para evaluar su desempeño y el área ocupada de ca-

¹Acrónimo del inglés “Reduced Instruction Set Computer”.

da una. Para la evaluación de estos diseños se deberá disponer de varios sumadores y multiplicadores que exploten el paralelismo presente en la evaluación de redes neuronales para el reconocimiento de imágenes, cuyas estructuras contienen cientos de miles de neuronas. Estos resultados podrían servir para definir funciones de similitud entre imágenes, es decir, dos imágenes que son clasificadas con el mismo peso por la red, se consideran muy similares.

Resultados y Objetivos

Los estudios realizados sobre el modelo de espacios métricos permitirán mejorar el desempeño de los MAMs analizados y estudiar la aplicación de los resultados obtenidos a otros [3, 4, 5, 9, 8].

Se profundizará el estudio del diseño de estructuras de datos, mejorando su eficiencia y adaptándolas mejor al nivel de la jerarquía de memorias donde se almacenarán y a las características de los datos a ser indexados. Se espera brindar nuevas herramientas eficientes de administración para bases de datos métricas, que logren acercar su desarrollo al de los modelos tradicionales de base de datos.

Se continuará analizando la manera de resolver consultas eficientemente, sin la utilización de índices. Además, se espera mejorar el desempeño de las operaciones de bajo nivel en los DBMS, mediante una nueva arquitectura del procesador.

Actividades de Formación

Se están formando investigadores en:

Doctorado en Cs. de la Computación: una tesis sobre expresividad de lenguajes lógicos de consulta.

Maestría en Cs. de la Computación: una tesis sobre búsqueda por similitud aproximada y otra sobre un índice dinámico eficiente.

Maestría en Informática: una tesis, de la Universidad Nacional de San Juan, sobre un índice dinámico para búsquedas aproximadas en disco.

Ingeniería en Computación: un trabajo de fin de carrera sobre diseño de sumadores/restadores y multiplicadores de punto flotante.

Referencias

- [1] G. Blaauw and F. Brooks, Jr. *Computer Architecture: Concepts and Evolution*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1997.
- [2] E. Chávez, K. Figueroa, and G. Navarro. Effective proximity retrieval by ordering permutations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(9):1647–1658, Sept 2008.
- [3] E. Chávez, M. Di Genaro, N. Reyes, and P. Roggero. Decomposability of disat for index dynamization. *Computer Science & Technology*, pages 110–116, 2017.
- [4] E. Chávez, V. Ludueña, N. Reyes, and F. Kasían. All near neighbor graph without searching. *Computer Science & Technology*, page 7, 2018. Por aparecer.
- [5] E. Chávez, V. Ludueña, N. Reyes, and P. Roggero. Faster proximity searching with the distal SAT. *Inf. Systems*, 59:15 – 47, 2016.
- [6] E. Chávez, G. Navarro, R. Baeza-Yates, and J. Marroquín. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273–321, September 2001.
- [7] P. Ciaccia and M. Patella. Approximate and probabilistic methods. *SIGSPATIAL Special*, 2(2):16–19, 2010.
- [8] K. Figueroa, C. Martínez, R. Paredes, N. Reyes, and P. Roggero. Dynamic list of clustered permutations on disk. In *Computer Science and Technology Series: XXI Argentine Congress of Computer Science Selected Papers*, pages 201–211. EDULP, 2016.
- [9] G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Inf. Systems*, 59:48 – 78, 2016.
- [10] R. Paredes, E. Chávez, K. Figueroa, and G. Navarro. Practical construction of k -nearest neighbor graphs in metric spaces. In *Proc. 5th Workshop on Efficient and Experimental Algorithms (WEA)*, LNCS 4007, pages 85–97, 2006.
- [11] H. Samet. *Foundations of Multidimensional and Metric Data Structures (The Morgan Kaufmann Series in Computer Graphics and Geometric Modeling)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2006.
- [12] P. Zezula, G. Amato, V. Dohnal, and M. Batko. *Similarity Search: The Metric Space Approach (Advances in Database Systems)*. Springer-Verlag New York, 2005. XVIII, 220 p., Hardcover ISBN: 0-387-29146-6.

BUENAS PRÁCTICAS EN EL DISEÑO DE ESTRUCTURAS DE DATOS EN BASES DE DATOS RELACIONALES

Muñoz Roberto Miguel, Maldonado Calixto, Damiano Luis Esteban, Romero María Soledad, Bueno Matías, Quinteros Sergio Ramón, Guevara Andrea, Peretti Juan Pablo, Carrasco Agustín, Urbano Barbara, Arguello Santiago

Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información (CIDS)
Departamento Ingeniería en Sistemas de Información
Facultad Regional Córdoba / Universidad Tecnológica Nacional
Maestro Marcelo López esq. Cruz Roja Argentina – Ciudad Universitaria - Córdoba
0351 - 4686385

{robertmunioz, calixtomaldonado, luis.damiano, romeroma.soledad, matiasbueno, ser.quinteros, andrezza77, peretti.juan, asermax, sbaarby}@gmail.com

RESUMEN

El propósito de este proyecto de investigación es desarrollar un conjunto de reglas de buenas prácticas para el diseño de modelos conceptuales para bases de datos relacionales.

El grupo de trabajo parte de una idea expresada en el libro Bases de Datos [1] y que en este proyecto fue transformada en hipótesis. La hipótesis de trabajo es que el diseño de un modelo conceptual de base de datos relacional, para un problema determinado y acotado, tiene una solución única, la cual puede ser validada en el resto del artículo.

Están planificadas las actividades y acciones que permitan recolectar información para analizar la problemática, con instrumentos tales como: encuestas, entrevistas personales y grupales y experimentos, incluyendo el análisis cuantitativo y cualitativo de los datos.

Con este proyecto de investigación se logrará como resultado final un documento que exprese los resultados de la investigación, que afirmen o nieguen la hipótesis. Es objetivo es obtener un compendio de reglas y principios de buenas prácticas, en el diseño del modelo conceptual de bases de datos relacionales. El grupo de investigación aspira a darle continuidad en el tiempo, publicarlo en medios especializados, incluirlo en capítulos en libros de la temática, darle visibilidad en reuniones científicas y

presentarlo públicamente en el ámbito académico.

Palabras clave: Bases de datos, modelo relacional, arquitecto de datos, estructura de datos, diseño de bases de datos.

CONTEXTO

El proyecto ha sido homologado por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional, con un tiempo de realización de 24 meses, desde el 1 de enero de 2018. La Unidad Científico-Tecnológica donde se desarrolla es el Centro de Investigación, Desarrollo y transferencia de Sistemas de Información (CIDS), de la Universidad Tecnológica Nacional, Facultad Regional Córdoba (UTN-FRC).

El origen del proyecto se centra en observaciones realizadas por los docentes de la cátedra Gestión de Datos, de la carrera de Ingeniería en Sistemas de Información, de la UTN-FRC. En este ámbito, y también en el desarrollo profesional de ellos, se pudo ver que los modelos conceptuales, solución de un problema de datos, tenían una mirada homogénea por parte de quienes discutían la solución, encontrando pocas o ninguna discrepancia.

Cabe aclarar que los docentes también se desempeñan en cátedras relacionadas a las siguientes temáticas: Programación, Bases de

Datos, Dirección de Proyectos e Integración de Sistemas de Información.

El equipo ha trabajado de manera continua en proyectos de investigación, relacionados a la temática de bases de datos, motores de bases de datos y específicamente en diseño de estructuras de datos:

- TecnoDB - Administrador de Base de Datos Relacional (2007) [2]
- PROMETEO - Desarrollo de un método y una herramienta para el aprovechamiento de Metadatos de Base de Datos Relacionales (2010) [3]
- Análisis y aplicación de metodologías para la generación de consultas complejas utilizando esquemas OLAP (2010) [4]
- Generador Automático de Modelos de Datos Normalizados en Bases de Datos Relacionales (2014-2015) [5]
- MultiDB. Plataforma Web para acceder a diferentes Sistemas de Gestión de Bases de Datos Relacionales (2016-2017) [6]

1. INTRODUCCIÓN

El modelo relacional es una forma de estructuración de datos, para su almacenamiento y manipulación. Abarca distintos estadios que van desde la comprensión de la solución de un problema de almacenamiento, pasando por el modelo ontológico de éste (habitualmente denominado modelo conceptual), hasta la implementación de esta solución a través de un motor de base de datos.

El arquitecto de la base de datos, para entender un problema de gestión y almacenamiento, debe en primer lugar poder interactuar con los expertos en el dominio del problema, que generalmente no son especialistas en la tarea de modelar estructuras de datos. Dichos expertos, en el análisis, diseño y dominio del problema, suelen ser uno o varios profesionales involucrados con el flujo de datos y sus características, sin ser necesariamente profesionales con conocimiento formal sobre el diseño de bases de datos. El arquitecto tiene que ser una persona con amplitud en su forma de

compresión, en la lectura de documentación relacionada al problema que está abordando, pues ésta puede ser fundamental para la identificación de aspectos procedimentales, restricciones técnicas y otros limitantes que puedan surgir de la misma.

En términos generales un arquitecto de base de datos debe poder aprender del problema que está tratando tanto como pueda, a fin de alcanzar una comprensión clara del dominio del problema, como expresa Giarratano en su libro *Sistemas Expertos...*[7], en página 4, “el dominio del conocimiento en el experto, debe crecer hasta comprender el dominio del problema”.

Los autores del libro *Base de Datos* [1], en la página 69 indicaron: “Es conveniente que se tenga una idea muy clara sobre el dominio del problema y cuál es su límite, pues esto permitirá que se alcance un resultado definido y limitado que se manejará con facilidad”. De esta manera, al especificar un límite concreto, claro y conceptualmente definido, se separa el estricto interés que tratará la solución del conocimiento general y extenso, de cualquier área a la que pertenezca el problema. En otros términos es una zona acotada, con un límite visible e invariable, que es una restricción del conocimiento general a su área de pertenencia. Es de vital importancia que el arquitecto conceptualice el modelo desde un entendimiento y comprensión, y no desde una interpretación. Esta última puede alejarlo de la realidad del problema tratado.

El arquitecto, habiendo recabado suficiente información del equipo de análisis y diseño sobre el problema que está tratando, será capaz de realizar su tarea con experticia, desarrollando el diseño de la estructura de datos (especificación de la conceptualización).

2. LÍNEA DE INVESTIGACIÓN, DESARROLLO

La conceptualización de la solución sobre un problema analizado, que se representa a través de un esquema (gráfico), compuesto por entidades, atributos y relaciones; habitualmente

se llama esquema conceptual o modelo conceptual.

Las entidades se identifican por un nombre, que es consecuencia de la reunión de un conjunto de atributos únicos en ésta, que la caracterizan y califican. Una entidad es única dentro de un modelo conceptual o estructura de datos resultado. Mantienen un estado de asociación entre sí, que están representados por la vinculación que se produce entre las entidades a través de las claves foráneas y primarias. Es imperativo que el estado de asociación esté correctamente planteado, pues es fundamental para asegurar la integridad de los datos en la estructura diseñada.

La clave primaria dentro de cada entidad está compuesta por un atributo, o un conjunto de estos, que producen la identificación de una tupla en forma unívoca e inequívoca dentro de la entidad. Esta clave primaria, que debe cumplir con las propiedades de unicidad y minimalidad [8], tiene una relación unívoca con el conjunto de atributos no clave que representa.

Para que una entidad quede relacionada a otra se debe establecer la vinculación, incorporando dentro de ésta los atributos de la clave primaria de la entidad con la que se desea vincular. De esta manera se identifica a un conjunto de atributos que recibe la denominación de clave foránea o ajena.

Definiendo las entidades con su clave primaria, y su estado de asociación con las claves foráneas, se obtiene el modelo conceptual - lógico del resultado terminado.

Un arquitecto de datos puede emplear cualquier procedimiento para determinar la estructura de datos, no hay uno definido. Suelen utilizarse metodologías: ascendente, denominada diseño por síntesis, o descendente, diseño por análisis [9]. Con la experiencia termina elaborando su proceso/receta.

Académicamente el grupo de docentes basa la enseñanza de estos temas en aplicar la técnica de normalización, con el análisis de dependencias funcionales y la aplicación de las reglas definidas como formas normales por

Codd [8], con un proceso “que sigue un método descendente evaluando cada relación contra el criterio de las formas normales y descomponiendo las relaciones según sea necesario” [9]. La estructura de datos lograda se menciona según el nivel alcanzado, por ejemplo: Normalización a tercera forma normal, normalización a segunda forma normal, etc. Este estado alcanzado por la solución, está vinculado a la entidad que alcanzó la forma normal menos avanzada de nivel dentro de la estructura resultado. Por ejemplo: para que una estructura de datos tenga la calificación de normalizada a tercera forma normal, todas las entidades componentes de la estructura deben estar en tercera forma normal. Si una de ellas se encuentra en segunda forma normal el resultado total está en segunda forma normal. Como ya se mencionó, la bibliografía avala y expresa los aspectos teóricos del diseño de bases de datos, con mayor o menor fluidez.

Como consecuencia de las prácticas en el aula y el ejercicio profesional de los integrantes, en el diseño de estructuras de datos, se ha observado que las posibilidades de desarrollar estructuras equivalentes (dos o más estructuras distintas) que proporcionen el mismo alcance de resultado no han sido demostradas hasta este momento.

En el año 2010 un grupo de docentes de la Universidad Tecnológica Nacional (Ing. Roberto Muñoz, Ing. Calixto Maldonado, Lic. Luis Damiano, Ing. Enrique Reinoso y el Ing. Maximiliano Abrutsky) escribieron un libro, para la editorial AlfaOmega, con el título: Base de Datos [1] y en él se presentó el pensamiento relacionado con esta afirmación que se acaba de esbozar. Este libro fue sometido, por parte de la editorial, a la consideración de distintos especialistas en bases de datos, de universidades locales y también de latinoamérica, y todos coincidieron que la afirmación no tenía precedente bibliográfico y admitieron aceptar el concepto con la siguiente salvedad en la página 67 [1]: “... Si bien el pasaje de un conjunto de relaciones no normalizadas a un conjunto de relaciones se

puede realizar de cualquier forma, el resultado debe ser siempre el mismo, independientemente del método que se haya utilizado. La bibliografía consultada en la elaboración de este capítulo, difiere con este autor en la manera de abordar el concepto “normalización”. Sin embargo, la metodología aquí utilizada se considera indispensable para la comprensión total del tema.”

Por todo esto, el grupo de investigación propone la discusión respecto a la siguiente hipótesis a demostrar “Si distintos arquitectos de bases de datos:

- a) parten del mismo origen de datos y los entienden como la misma colección de datos,
- b) observan un problema de almacenamiento de datos inicial, por ejemplo, una planilla que contiene el conjunto de datos (una estructura sin el análisis suficiente),
- c) poseen el mismo nivel de conocimiento acerca del dominio del problema, contando con la misma documentación explicativa, y
- d) todos pueden ser considerados como expertos, sobre cómo realizar la transformación desde un conjunto de datos hasta llegar a una estructura de datos deseable;

entonces los arquitectos pueden realizar la transformación con cualquier método de razonamiento o con cualquier procedimiento lógico, pero todos siempre deben llegar al mismo resultado”.

Esto es lo que permitiría validar o no la hipótesis de trabajo.

3. RESULTADOS ESPERADOS

El proyecto formalmente inició en enero del corriente año, por lo cual no se dispone aún de resultados, pero ya se han enunciado líneas de trabajo y un cronograma a seguir para conseguir los avances previstos.

Es objetivo, del grupo de investigación, proponer un conjunto de buenas prácticas para el diseño del modelo conceptual de datos

normalizado, para bases de datos relacionales, tendientes a explicar las reglas y características que fundamentan que el modelo conceptual de una base de datos relacional puede ser único, para un dominio dado.

Los resultados esperados del proyecto se pretenden documentar y de esta manera lograr los siguientes objetivos:

- Determinar los principios que argumentan la hipótesis.
- Contrastar y ejemplificar estructuras validadas de modelos conceptuales con posibles estructuras no válidas.
- Ejemplificar prácticamente la hipótesis.
- Identificar las premisas teóricas y prácticas que sustentan la hipótesis.
- Identificar todas las estructuras típicas que se utilizan en universo de soluciones.
- Identificar las estructuras diferentes que ofrecen una solución equivalente, si las hay, mostrando las similitudes y discrepancias de esas estructuras.
- Explicar la forma en que implementan una estructura de datos válida los arquitectos de datos.
- Sugerir un método de trabajo que permita lograr una estructura deseable, en la búsqueda de estandarizar el proceso.
- Documentar la estandarización de las acciones identificadas, junto a las reglas.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está constituido por: a) un director de proyecto, b) un codirector, c) cinco docentes investigadores, dos en formación y tres de apoyo, d) dos estudiantes becarios y e) un graduado.

Todos los docentes, que integran el equipo de investigación, comparten su actividad académica en la cátedra de Gestión de Datos, de tercer nivel de la Carrera de Ingeniería en Sistemas de Información, donde es central la temática de Bases de Datos. En el contenido mínimo de la asignatura el diseño de una estructura de datos ocupa un espacio

preponderante y en el tiempo de desarrollo ocupa la primera mitad de un cuatrimestre.

El reciente graduado, Ing. en Sistemas de Información Agustín Carrasco, invitado a participar en el proyecto, ha sido ya integrante de anteriores Proyectos de Investigación y como estudiante logró ganar en dos ediciones consecutivas el Concurso de Programación que se realiza en la UTN-FRC.

Los estudiantes Bárbara Romina Urbano Moreno y Santiago Manuel Argüello están cursando actualmente el cuarto año de la Carrera de Ingeniería en Sistemas de Información e iniciando el camino en investigación como becarios.

5. BIBLIOGRAFÍA

[1] Reinoso, E.; Maldonado, C.; Muñoz, R.; Damiano, L.; Abrutsky, M. - Bases de Datos - Edit. AlfaOmega Editores - Argentina - 2012- ISBN: 978-987-1609-31-4.

[2] Gastañaga, I., Maldonado, C., Martínez Spessot, C. I., & Hintermeister, E. (2006). TecnoDB una Base de Datos Relacional y Prometeo un método de aprovechamiento de Metadatos y Generador de Consultas. In VIII WICC.

[3] Marciszack, M., Maldonado, C., Martínez Spessot, C. I., Muñoz, R., Navarro, A., Peretti, J. P., & Roggero, L. (2009). Prometeo: una herramienta para el aprovechamiento de metadatos de base de datos relacionales. In XI WICC.

[4] Vaca, P. A., & Abrutsky, M. (2009). Herramienta para consultas complejas orientada a usuarios finales. In XI WICC.

[5] Paz Menvielle, M. A., Cuevas, J. C., Damiano, L. E., Muñoz, R., & Quinteros, S. (2013, June). Generador automático de modelos de datos normalizados en bases de datos relacionales. In XV WICC.

[6] Muñoz Roberto Miguel, Maldonado Calixto, Damiano Luis Esteban, Romero María Soledad, Cuevas Juan Carlos, Quinteros Sergio Ramón, Guevara Andrea, Carrasco Agustín. MultiDB, Plataforma Web para acceder a

diferentes Sistemas de Gestión de Bases de Datos Relacionales. XVIII WICC (2016)

[7] Giarratano, Joseph - Riley, Gary - Sistemas Expertos Principios y Programación - Editorial Cengage Learning / Thomson Internacional, 2005 - ISBN 9789706860590

[8] Date, C.J. Introducción a los Sistemas de Bases de Datos. Volumen 1. Quinta Edición. USA, 1993. Edit. Addison Wesley Iberoamericana, Inc. ISBN: 0-201-51859-7.

[9] Elmasri, R. y Navathe, S. Fundamentos de Sistemas de Bases de Datos. 5ta Edición, 2007 Edit. Pearson. ISBN: 978-84-7829-085-7.

Recuperación de Datos e Información en Bases de Datos Masivas

Luis Britos, Fernando Kasián, Verónica Ludueña, Franco Merenda,
Marcela Printista, Nora Reyes, Patricia Roggero

LIDIC, Dpto. de Informática, Fac. de Cs. Físico Matemáticas y Naturales, Universidad Nacional de San Luis
{lebritos, fkasian, vlud, mprinti, nreyes, proggero}@unsl.edu.ar, merenda.franco83@gmail.com

Edgar Chávez

Centro de Investigación Científica y de Educación Superior de Ensenada, México

elchavez@cicese.mx

Claudia Deco

Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario

deco@fceia.unr.edu.ar

Resumen

En la actualidad es cada vez más evidente la necesidad de procesar grandes conjuntos de datos, de manera tal de poder obtener información útil a partir de ellos. Sin embargo, la evolución de las tecnologías de información y comunicación, en conjunto con la gran cantidad y variedad de información disponible digitalmente, han llevado en las últimas décadas al surgimiento de nuevos depósitos no estructurados de información, en los cuales los datos que no se adaptan fácilmente al modelo relacional. A tipos de datos tales como texto libre, imágenes, audio, video, secuencias biológicas de ADN o proteínas, entre otros; no se los puede estructurar más en claves y registros, o tal estructuración es muy dificultosa (tanto manual como computacionalmente), y restringe de antemano los tipos de consultas que luego se pueden realizar. Como muchas aplicaciones computacionales necesitan recuperar datos e información desde estas grandes bases de datos conteniendo datos no estructurados, es necesario lograr eficiencia en formas más sofisticadas de búsqueda que la habitual sobre datos estructurados. Así, dada una consulta, el objetivo de un sistema de recuperación de información es obtener lo que podría ser útil o relevante para el usuario, usando una estructura de almacenamiento especialmente diseñada para responderla eficientemente.

Palabras Claves: bases de datos masivas, computación de alto desempeño, recuperación de información.

1. Contexto

Esta línea de investigación se encuentra enmarcada dentro del Proyecto Consolidado 3-30114 de la Universidad Nacional de San Luis (UNSL) y en el Programa de Incentivos (Código 22/F434): “Tecnologías Avanzadas Aplicadas al Procesamiento de Datos Masivos”, dentro de la línea “Recuperación de Datos e Información”, desarrollada en el Labo-

ratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC) de la UNSL. Actualmente se encuentra en proceso de evaluación la nueva presentación.

En esta línea de investigación se busca desarrollar herramientas eficientes para sistemas de información sobre bases de datos masivas, conteniendo datos no estructurados. Por lo tanto, se analizan nuevas técnicas que permitan una buena interacción con el usuario, nuevas estructuras de datos (índices) capaces de manipular eficientemente datos no estructurados y que puedan utilizarse para administrar bases de datos masivas que contienen datos no estructurados. Así, se pretende contribuir a la incorporación de información no estructurada en los procesos de toma de decisiones y resolución de problemas, no considerados en los enfoques clásicos. Por lo tanto, el objetivo principal de esta línea es el diseño y desarrollo de índices que sirvan de apoyo a sistemas de recuperación dedicados a conjuntos de datos no estructurados masivos tales como: datos multimedia, texto, secuencias de ADN, etc. , permitiendo que estos sistemas cuenten con estructuras de datos eficientes y escalables, para memorias jerárquicas, que hagan uso, de ser necesario, de técnicas de computación de alto desempeño (HPC).

2. Introducción y Motivación

En la actualidad, gracias al uso masivo de internet, se ha producido una significativa aceleración tanto en el crecimiento del volumen de datos capturados y almacenados, como en la creciente variación en los tipos de datos que aparecen. En este contexto, se hace necesario que las técnicas tradicionales para el

procesamiento, análisis y obtención de información útil deban ser redefinidas para formular nuevas metodologías de abordaje.

En general, los sistemas tradicionales de computación utilizan principalmente información estructurada, la cual puede organizarse en claves y registros, sobre los cuales tiene sentido aplicar búsquedas tradicionales y donde su estructura puede interpretarse y utilizarse en programas casi directamente. Pero, por el volumen y variedad de los datos disponibles actualmente, dos de las características de los datos en el ámbito de problemas de “big data”, no es posible restringirse a búsquedas sobre datos estructurados, porque obligaría a representar una visión parcial del problema, dejando fuera información que podría ser relevante para la resolución efectiva del mismo. Por lo tanto, en la era de “big data” es necesario administrar eficientemente información no estructurada y considerar tipos de búsqueda más generales que puedan servir de apoyo, por ejemplo, en la toma de decisiones. Uno de estos tipos de búsqueda más generales son las búsquedas por similitud, las cuales se suelen sustentar sobre métodos de acceso o índices métricos [3], que permiten responderlas más eficientemente.

Así, dada la gran cantidad de datos con los que se trabaja, ante consultas de recuperación de información sobre bases de datos conteniendo datos no estructurados, se pueden utilizar estos índices para lograr eficiencia en la respuesta. Dichos índices pueden tener distintas características que los hacen indicados para aplicaciones reales: eficientes, dinámicos, escalables, resistentes a la *maldición de la dimensión*, entre otras. Un enfoque útil para sistemas de recuperación usando búsqueda por similitud es “la búsqueda basada en contenidos”, la cual usa el dato no estructurado en sí mismo para describir lo que se busca. Para calcular la similitud entre dos objetos, se debe definir una función de distancia que permita describir realmente la disimilitud entre ellos.

El modelo habitual para las búsquedas por similitud es el de espacios métricos; dado que, además de brindar un marco formal, es independiente del dominio de la aplicación. Un espacio métrico se compone de un universo \mathcal{U} de objetos y una función de distancia $d : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}^+$, la cual cumple con las propiedades de una métrica. Sobre una *base de datos* $\mathcal{S} \subseteq \mathcal{U}$, se pueden considerar dos tipos de búsqueda por similitud: la *búsqueda por rango* y la *búsqueda de los k vecinos más cercanos*. La función de distan-

cia permite medir el mínimo esfuerzo (costo) necesario que se debe realizar para transformar un objeto en otro. Dependiendo de los tipos de datos no estructurados, el cálculo de la distancia puede ser muy costoso. Por lo tanto, se busca ahorrar cálculos de distancia y ello se logra generalmente gracias a que la función de distancia cumple con la desigualdad triangular.

Si se considera que la base de datos \mathcal{S} posee n objetos, trivialmente cualquier consulta se puede responder con n evaluaciones de distancia. Sin embargo, en la mayoría de las aplicaciones sobre conjuntos de datos masivos, como las distancias son costosas de computar (por ej.: comparación de huellas digitales), no es factible aplicar la solución trivial. Así, para responder consultas con la menor cantidad de cálculos de distancia posibles se debe preprocesar la base de datos para construir un índice. En algunos casos, es probable que la base de datos, el índice, o ambos, no puedan almacenarse en memoria principal. Por lo tanto, para lograr eficiencia, se debe minimizar el número de operaciones de E/S, considerar la jerarquía de memorias, en algunos casos admitir respuestas no exactas y cuando sea posible utilizar técnicas paralelas.

Así, en este contexto se considera como objetivo obtener herramientas de recuperación de información, desarrollando nuevas técnicas y aplicaciones que soporten la interacción con el usuario, diseñando índices capaces de manipular eficientemente grandes volúmenes de datos no estructurados y facilitando la realización de diferentes tipos de consultas, de modo de contribuir al desarrollo de aplicaciones reales para problemas de big data.

3. Líneas de Investigación

Como se pretende investigar sobre distintos aspectos de los sistemas de recuperación de información sobre grandes volúmenes de datos no estructurados, se ha considerado el diseño de nuevos índices y la resolución de distintas consultas sobre estos tipos de bases de datos y cómo lograr eficiencia y escalabilidad en las soluciones al considerar grandes volúmenes de datos.

Índices

Los índices que resultan apropiados, para luego realizar búsquedas sobre bases de datos conteniendo datos no estructurados, son los índices métricos [3]. En todos ellos se aprovecha que la función de distancia debe cumplir la propiedad de desigualdad

triangular para ahorrar algunos cálculos de distancia y de esta manera tiempo, gracias a que la desigualdad triangular permite estimar la distancia entre cualquier objeto de consulta q y los objetos de la base de datos, si se mantienen algunas distancias precalculadas entre los elementos de la base de datos y elementos distinguidos. Los dos enfoques más comunes se diferencian en si esos objetos distinguidos son *pivotes* o *centros*. Si son pivotes se almacenan las distancias de todos los objetos de la base de datos a ellos y si por el contrario son centros se particiona el espacio en zonas denominadas *particiones compactas*, por cercanía a los centros y se almacena un radio de cobertura para determinar la zona de cada centro.

En nuestro caso, nos enfocamos en diseñar buenos índices que consideren:

Dinamismo: Los índices pueden construirse de manera estática, si los objetos de la base de datos se conocen de antemano. En estos índices denominados *estáticos* las búsquedas se realizan luego de construido el índice. Por el contrario, si no se pueden tener los objetos de antemano y la única manera de construir el índice es a medida que se incorporan los elementos; es decir, de manera incremental, se considera que las búsquedas pueden realizarse en cualquier momento. Esta clase de índices se denominan *dinámicos*. Los índices estáticos, por conocer a toda la base de datos, pueden seleccionar los mejores objetos distinguidos para una estructura de datos determinada. En cambio, en los índices dinámicos esto no es posible.

Jerarquía de Memorias: Otro aspecto importante para buscar una solución es saber si se puede trabajar en memoria principal o, por el contrario, si por ser conjuntos de datos masivos se deberá trabajar en otros niveles de la jerarquía de memorias. En caso que el índice deba alojarse en memoria secundaria, se deben minimizar la cantidad de cálculos de distancia y también el número de operaciones de E/S.

Computación de Alto Desempeño: En algunos casos, si no se logra la eficiencia deseada mediante la optimización del índice en sí mismo, se pueden aplicar técnicas de computación de alto desempeño (HPC) para acelerar los tiempos de respuesta a las consultas.

Exactitud de la Respuesta: Otra manera de acelerar la respuesta a una consulta por similitud es admitir una respuesta aproximada, permitiendo que la

misma sea de menor calidad o menos exacta, pero muy rápida.

Dimensionalidad Intrínseca: los índices para búsquedas por similitud, al trabajar sobre el modelo de espacios métricos, pueden también sufrir de la llamada *maldición de la dimensión* [3]; es decir, los índices se degradan a medida que la dimensión de los espacios aumenta. Existen índices que se comportan mejor en espacios difíciles (dimensión intrínseca mediana a alta) y otros que son adecuados para espacios fáciles (dimensión intrínseca baja).

Como nuestro interés está puesto sobre conjuntos de datos masivos que contienen datos no estructurados, los volúmenes de información con los que se debe trabajar (por ejemplo, millones de imágenes en la Web) hace necesario que los índices sean almacenados en memoria secundaria. En este caso, para lograr eficiencia, no sólo se debe considerar que las búsquedas realicen el menor número de cálculos de distancia sino también, dado el costo de las operaciones de E/S, se efectúe la menor cantidad posible de operaciones sobre el disco. Por ello, esta línea se dedica a diseñar índices especialmente adaptados para trabajar en memoria secundaria, cuyo desempeño en las búsquedas sea bueno. Así, se ha diseñado e implementado una versión paralela del *Conjunto Dinámico de Clusters* (DSC) [8]. Este índice, basado en la *Lista de Clusters* (LC) [2], está especialmente diseñado para memoria secundaria y es completamente dinámico, admite inserciones y eliminaciones y tiene un buen desempeño en las búsquedas, principalmente en la cantidad de operaciones de E/S. DSC ha demostrado ser muy competitivo frente a otras de las buenas estructuras del estado del arte. Por lo tanto, se buscará aplicar y comparar distintas estrategias de paralelización con el fin de determinar la más adecuada.

El *Árbol de Aproximación Espacial Distal* (DiSAT), basado en el *Árbol de Aproximación Espacial* [6], es un índice estático que no necesita sintonizar ningún parámetro y es muy eficiente gracias a definir una partición de hiperplanos con muy buenas características [4]. La raíz elegida para el DiSAT define una partición sobre el espacio, donde las zonas que se obtienen son muy compactas y los hiperplanos que las definen permiten diferenciarlas muy bien. Por ello, se busca aprovechar la información que brindan distintas particiones sobre el espacio para clasificar los elementos de acuerdo a las zonas en las que cada elemento cae en las distintas par-

ticiones consideradas. En este caso, a cada elemento se le asigna una secuencia de bits, denominada “sketch”, donde cada bit indica de qué lado del hiperplano considerado se encuentra el elemento. Este conjunto de “sketches” constituye el índice en sí mismo. Cuando se considera una consulta, se calcula el sketch del elemento de consulta q y se lo compara con los sketches de todos los elementos de la base de datos, sin calcular realmente distancias entre objetos sino entre sketches y se revisan luego los objetos más prometedores primeros. En este caso, se espera que un elemento similar a q estará en una partición similar en el espacio. En este caso, se puede limitar de antemano el número de distancias reales que se permiten calcular, logrando una respuesta aproximada a la consulta por similitud con poco costo.

Existen en la actualidad pocas medidas que permitan reflejar adecuadamente la dimensionalidad intrínseca de los espacios métricos [3]. Sin embargo, si se pudiera calcular la dimensionalidad intrínseca de un espacio métrico con cierta confiabilidad, se podría elegir el índice que tuviera mejor desempeño en las búsquedas para esa dimensión en particular. Por lo tanto, se han propuesto nuevas medidas de evaluación de la dimensionalidad intrínseca y se las ha evaluado experimentalmente junto a otras medidas ya conocidas, para ver cuál de ellas puede reflejar de manera más confiable la dimensionalidad intrínseca de un espacio métrico [7].

Por otra parte, se está estudiando cómo aprovechar los índices para búsquedas por similitud sobre bases de datos masivos de datos no estructurados, para solucionar un problema de estacionamiento de vehículos, usando en este caso los índices como herramienta de apoyo en un sistema de recuperación de datos e información.

En esta línea de investigación se están desarrollando dos tesis de maestría y un trabajo final.

Sistema Administrador para Bases de Datos Multimedia

A pesar de que las operaciones más comunes sobre bases de datos multimedia son las búsquedas por rango o de k -vecinos más cercanos, existen otras operaciones de interés tales como las distintas variantes del *join* por similitud. La operación de *join* por similitud se considera una de las operaciones que debería brindar típicamente un sistema administrador para bases de datos multimedia [10].

Existen diferentes variantes para el *join* por similitud, dependiendo del criterio de similitud utilizado, pero ellas tienen en común que se aplican entre

dos bases de datos A y B , ambos subconjuntos del mismo universo del espacio métrico \mathcal{U} que modela a la base de datos multimedia. El resultado de cualquiera de las variantes del *join* por similitud entre A y B obtendrá el conjunto de pares formados por un objeto de A y otro de B , tales que entre ellos se satisface el criterio de similitud considerado. Las variantes más conocidas son: el *join* por rango, el *join* de k -vecinos más cercanos y el *join* de k pares de vecinos más cercanos; entre otras.

Formalmente, dadas $A, B \subseteq \mathcal{U}$, se define el *join por similitud* entre A y B ($A \bowtie_{\Phi} B$) como el conjunto de todos los pares (x, y) , donde $x \in A$ e $y \in B$; es decir, $(x, y) \in A \times B$, tal que (x, y) es verdadero (se satisface el criterio de similitud entre x e y). Al resolver el *join* por similitud es posible que ambas, una o ninguna de la bases de datos posean un índice; o que ambas bases de datos se indexen conjuntamente con un índice diseñado para el *join*. Calcular cualquiera de las variantes del *join* por similitud de manera exacta es muy costoso [9], así vale la pena analizar posibilidades de obtener una respuesta aproximada al *join*, más rápidamente, aunque siempre buscando buena calidad en la respuesta.

PostgreSQL es el primer sistema de base de datos que permite realizar consultas por similitud sobre algunos atributos, particularmente indexa para búsquedas de k -vecinos más cercanos (índices *KNN-GiST*). Estos índices pueden ser usados sobre texto, comparación de ubicación geoespacial, etc. Sin embargo, los índices *K-NN GiST* proveen plantillas sólo para índices con estructura de *árbol balanceado* (*B-tree*, *R-tree*), pero el “balance” no siempre es bueno para los índices que se utilizan en búsquedas por similitud [1]. Además, no se dispone de este tipo de consultas para todo tipo de datos métricos. Así, es importante proveer un DBMS para bases de datos métricas que maneje todos los posibles datos métricos y las operaciones de interés sobre ellos [5].

Más aún, dado que las respuestas a consultas de *join* suelen ser conjuntos muy grandes de pares de objetos y muchos de esos pares son muy similares entre sí, se planea introducir sobre las operaciones de *join* la posibilidad de diversificar las respuestas [11]; es decir, un operador de *join* por similitud que asegure un conjunto más pequeño, más diversificado de respuestas útiles y, de ser posible, más rápido de obtener. Estos desarrollos, entre otros, permitirán tener un DBMS con mayores posibilidades de aplicación en sistemas de información reales.

Esta línea corresponde a una tesis de maestría.

4. Resultados

Se ha publicado la evaluación experimental de un conjunto de estimadores de la dimensión intrínseca de un espacio métrico [7], que permitió establecer que los mejores estimadores de dimensión son el exponente de distancia y el estimador basado en correlación.

Actualmente se está evaluando experimentalmente la versión paralela del índice *DSC*, que trabaja con grandes volúmenes de datos, diseñada especialmente para memoria secundaria, que admite inserciones y eliminaciones de elementos y que permitirá responder eficientemente a lotes de consultas por similitud. Además, se encuentra también en proceso de evaluación la propuesta de sketches basados en el *DiSAT*. Se continúa trabajando en la extensión de *PostgreSQL* para que brinde facilidades de soporte a más tipos de consultas por similitud, sobre distintos tipos de datos y que considere opciones de respuesta aproximada, como así también la posibilidad de obtener una respuesta diversificada en el caso de los joins por similitud.

5. Formación de Recursos

En esta línea se están realizando las siguientes tesis de Maestría en Ciencias de la Computación:

1. “Estructuras Eficientes sobre Datos Masivos para Búsquedas en Espacios Métricos”,
2. “Cómputo Aproximado del Grafo de Todos los k -Vecinos”,
3. “Sistema Administrador para Bases de Datos Métricas”.

Además, está en su etapa inicial el desarrollo de un trabajo final de la Ingeniería en Computación.

Referencias

- [1] E. Chávez, V. Ludueña, and N. Reyes. Revisiting the VP-forest: Unbalance to improve the performance. In *Proc. de las JCC08*, page 26, 2008.
- [2] E. Chávez and G. Navarro. A compact space decomposition for effective metric indexing. *Pattern Recognition Letters*, 26(9):1363–1376, 2005.
- [3] E. Chávez, G. Navarro, R. Baeza-Yates, and J. Marroquín. Searching in metric spaces. *ACM*, 33(3):273–321, September 2001.
- [4] Edgar Chávez, Verónica Ludueña, Nora Reyes, and Patricia Roggero. Faster proximity searching with the distal {SAT}. *Information Systems*, pages–, 2016. In Press, Available online.
- [5] F. Kasián and N. Reyes. Búsquedas por similitud en PostgreSQL. In *Actas del XVIII Congreso Argentino de Ciencias de la Computación (CACiC)*, pages 1098–1107, Bahía Blanca, Argentina, October 2012. Universidad Nacional del Sur.
- [6] G. Navarro. Searching in metric spaces by spatial approximation. *VLDBJ*, 11(1):28–46, 2002.
- [7] Gonzalo Navarro, Rodrigo Paredes, Nora Reyes, and Cristian Bustos. An empirical evaluation of intrinsic dimension estimators. *Information Systems*, 64:206 – 218, 2017.
- [8] Gonzalo Navarro and Nora Reyes. New dynamic metric indices for secondary memory. *Information Systems*, 59:48 – 78, 2016.
- [9] R. Paredes and N. Reyes. Solving similarity joins and range queries in metric spaces with the list of twin clusters. *JDA*, 7:18–35, March 2009. doi:10.1016/j.jda.2008.09.012.
- [10] C. Rong, C. Lin, Y. N. Silva, J. Wang, W. Lu, and X. Du. Fast and scalable distributed set similarity joins for big data analytics. In *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, pages 1059–1070, April 2017.
- [11] Lucio F. D. Santos, Luiz Olmes Carvalho, Willian D. Oliveira, Agma J.M. Traina, and Jr. Traina, Caetano. Diversity in similarity joins. In Giuseppe Amato, Richard Connor, Fabrizio Falchi, and Claudio Gennaro, editors, *Similarity Search and Applications*, volume 9371 of *Lecture Notes in Computer Science*, pages 42–53. Springer International Publishing, 2015.

Minería de Datos y Big Data. Aplicaciones en riesgo crediticio, salud y análisis de mercado

L. Lanzarini¹, W. Hasperué¹, A. Villa Monte^{1,3}, M. J. Basgall^{1,4}, R. Molina^{1,5}, L. Rojas Flores⁶, J. Corvi²,
P. Jimbo Santana⁷, A. Fernandez Bariviera⁸, C. Puente⁹, J. A. Olivas¹⁰

¹ Instituto de Investigación en Informática LIDI*, Facultad de Informática, UNLP, La Plata, Argentina *

² Facultad de Informática, Universidad Nacional de La Plata, La Plata, Argentina

³ Becario postgrado UNLP ⁴ UNLP, CONICET, III-LIDI, La Plata, Argentina ⁵ Becario CIN

⁶ Facultad de Ingeniería, Universidad Nacional de la Patagonia San Juan Bosco, Ushuaia, Argentina

⁷ Facultad de Ciencias Administrativas, Universidad Central del Ecuador, Quito, Ecuador

⁸ Dpto de Economía, Universitat Rovira i Virgili, Reus, España

⁹ Escuela Técnica Superior de Ingeniería ICAI, Universidad Pontificia Comillas, Madrid, España

¹⁰ Dpto. Tecnología y Sistemas de la Información, Universidad de Castilla-La Mancha, Ciudad Real, España

* Centro asociado de la Comisión de Investigaciones Científicas de la Pcia. De Bs. As. (CIC)

{laural, whasperue, avillamonte, mjbasgall}@lidi.info.unlp.edu.ar

{luisf.09, julieta.corvi}@gmail.com, pjimbo@pepsolutions.com, aurelio.fernandez@urv.net,
cristina.puente@icai.comillas.edu, joseangel.olivas@uclm.es

CONTEXTO

Esta presentación corresponde al proyecto “Sistemas inteligentes. Aplicaciones en reconocimiento de patrones, minería de datos y big data” (Periodo 2018–2021) del Instituto de Investigación en Informática LIDI.

RESUMEN

Esta línea de investigación se centra en el estudio y desarrollo de Sistemas Inteligentes para la resolución de problemas de Minería de Datos y Big Data utilizando técnicas de Aprendizaje Automático. Los sistemas desarrollados se aplican particularmente al procesamiento de textos y reconocimiento de patrones en imágenes.

En el área de la Minería de Datos se está trabajando, por un lado, en la generación de un modelo de fácil interpretación a partir de la extracción de reglas de clasificación que permita justificar la toma de decisiones y, por otro lado, en el desarrollo de nuevas estrategias para tratar grandes volúmenes de datos.

Con respecto al área de Big Data se están realizando diversos aportes usando el framework Spark Streaming. En esta dirección, se está investigando en una técnica

de clustering dinámico que se ejecuta de manera distribuida. Además se ha implementado en Spark Streaming una aplicación que calcula el índice de Hertz de manera online, actualizándolo cada pocos segundos con el objetivo de estudiar un cierto mercado de negocios.

En el área de la Minería de Textos se han desarrollado estrategias para resumir documentos a través de la extracción utilizando métricas de selección y técnicas de optimización de los párrafos más representativos. Además se han desarrollado métodos capaces de determinar la subjetividad de oraciones escritas en español.

Palabras clave: Minería de Datos, Minería de Textos, Big Data, Redes neuronales, Resúmenes extractivos, Sentencias causales temporales.

1. INTRODUCCION

El Instituto de Investigación en Informática LIDI tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos tipos de estrategias adaptativas. Los resultados obtenidos han sido medidos en la solución de problemas pertenecientes a distintas áreas. A continuación se detallan los resultados obtenidos durante el último año.

1.1. MINERÍA DE DATOS

Obtención de Reglas de Clasificación

Esta línea de investigación está centrada en el diseño de nuevos algoritmos para la obtención de conjuntos de reglas de clasificación con tres características principales: precisión adecuada, baja cardinalidad y facilidad de interpretación. Esto último está dado por el uso de un número reducido de atributos en la conformación del antecedente que, sumada a la baja cardinalidad del conjunto de reglas, permite distinguir patrones sumamente útiles a la hora de comprender las relaciones entre los datos y tomar decisiones [1]. La aplicación de estos métodos en la predicción de riesgo crediticio ha arrojado resultados satisfactorios [2,3].

Actualmente se está trabajando en la fuzificación de las reglas con el objetivo de facilitar aún más su comprensión por parte del agente que debe decidir el otorgamiento del crédito. Se ha comprobado que con sólo fuzificar el antecedente de la regla se obtienen conjuntos de reglas de clasificación con un incremento significativo en la precisión en relación a lo publicado en [4].

A futuro se incorporará, a la recomendación dada por la regla, un factor de confianza que ayude a discernir entre posibles recomendaciones. Este es un aspecto importante ya que además de las características propias del solicitante del crédito existen condiciones macroeconómicas que condicionan la respuesta.

1.2. BIG DATA

Aplicaciones en Big Data

En esta línea se trabaja sobre el procesamiento en streaming y en batch de grandes volúmenes de datos en formato texto. Para esto se están desarrollando estrategias que aplican técnicas de machine learning que presenten la característica de ser iterativas, operando sobre el conjunto completo de los datos de un flujo, brindando resultados en tiempos de respuestas cortos los cuales se

adaptan de manera dinámica a la llegada de nuevos datos [5, 6].

Estas técnicas dinámicas se están implementando en el framework Spark Streaming, adecuado para procesamiento paralelo, distribuido y online. En este framework se desarrolló una aplicación que permite el cálculo del índice de Hertz de manera online y dinámica, esto es, cada cierto tiempo la aplicación usa los nuevos datos recolectados y los procesa con aquellos que habían sido procesados previamente para poder hacer un seguimiento online de un cierto mercado de negocios [7].

Los temas que se abordan en esta línea abarcan la implementación de técnicas de clustering para el tratamiento de flujos de datos, la detección de tópicos, el análisis de sentimiento y el procesamiento de datos relacionados al comercio realizado con criptomonedas [8].

1.3. MINERIA DE TEXTOS

Hoy en día, la información que nos rodea lo hace en su gran mayoría en forma de texto. El volumen de información no estructurada crece continuamente de tal manera que resulta necesario separar por medio de técnicas de procesamiento de texto lo esencial de lo que no lo es así como distinguir proposiciones subjetivas de las objetivas.

Resumen Automático de Documentos

Esta línea de investigación se centra en la generación automática de resúmenes. Entre los enfoques existentes se ha puesto el énfasis en el extractivo cuyo resumen está formado por un subconjunto de sentencias de un documento seleccionadas apropiadamente. Actualmente, a partir del trabajo realizado en [9] se están analizando en la construcción de distintos tipos de resúmenes (1) el impacto de varias tareas de preprocesamiento de textos, (2) la participación de un conjunto amplio de métricas [10] y (3) la incorporación de semánticas en el análisis [11]. Para llevar a cabo estos experimentos se desarrolló una

herramienta de manipulación de documentos científicos programada en Python con MySQL utilizando las librerías NLTK, urllib y bs4, entre otras. Los experimentos están siendo realizados sobre artículos científicos publicados en PLOS ONE hasta tanto se consiga el acceso a las colecciones DUC.

Por otro lado, en [12] se estudió la relación entre algunos tipos de resúmenes extractivos y los formados únicamente por las sentencias causales detectadas en un documento. Este tipo de sentencias son de suma utilidad para analizar documentos clínicos por ser una componente principal de toda explicación médica. Ellas expresan, por ejemplo, las causas de las enfermedades o muestran los efectos de cada tratamiento. Actualmente, se están investigando las restricciones temporales asociadas a relaciones causales.

Clasificación de oraciones

Con el objetivo de analizar la subjetividad u objetividad de un texto se desarrolló una representación de oraciones escritas en español en formato vectorial que permite etiquetarlas. Esta representación utiliza distintas métricas lingüísticas para convertir una oración a una matriz numérica. Dado que la cantidad de filas de estas matrices depende de la longitud de la oración se realiza una normalización que convierte dicha matriz en un vector de longitud fija para poder comparar los vectores de distintas oraciones. Se han utilizado las redes neuronales y las máquinas de soporte vectorial para entrenar modelos que permitan clasificar una oración en objetiva o subjetiva. [13]

2. TEMAS DE INVESTIGACIÓN Y DESARROLLO

- Estudio de técnicas de optimización poblaciones y redes neuronales artificiales para la obtención de reglas difusas de tipo IF-THEN.
- Métodos estructurados y no estructurados a la representación de documentos.

- Problemas de clasificación con desbalance de clase severo.
- Representación de documentos de texto utilizando métricas.
- Obtención de resúmenes automáticos de texto.
- Implementación de técnicas en el paradigma de MapReduce
- Implementación del índice de Hertz en Spark streaming.
- Implementación de un algoritmo de clustering dinámico en Spark streaming.
- Propuesta de una representación vectorial de oraciones de longitud variable.
- Desarrollo de modelos que permiten clasificar oraciones en subjetivas u objetivas.

3. RESULTADOS OBTENIDOS

- Desarrollo de un método de obtención de reglas de clasificación difusas con énfasis en la reducción de la complejidad del modelo aplicable a riesgo crediticio.
- Desarrollo de una representación de términos que junto con un modelo de clasificación permite identificar palabras clave en un documento.
- Desarrollo de un algoritmo de clustering que selecciona el número de clusters de manera dinámica implementado en el frameworks Spark streaming.
- Implementación en Spark Streaming de una aplicación que calcula de manera online el coeficiente de Hurst y lo actualiza cada un cierto tiempo.
- Identificación de las partes relevantes de un documento. Propuesta de distintas métricas y una representación vectorial de oraciones de diferentes longitudes.
- Análisis y comparación de resúmenes extractivos de documentos.
- Implementación de modelos usando redes neuronales para la determinación de

subjetividad en oraciones extraídas en textos escritos en español.

- Aplicación de las sentencias causales en el desarrollo de un sistema que asista en la administración de medicamentos mediante el control de intervalos de tiempo.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 2 profesores con dedicación exclusiva, 1 becario doctoral UNLP, 1 becario doctoral CONICET, 1 becario CIN, 1 doctorando, 2 tesistas de grado y 3 profesores extranjeros.

Dentro de los temas involucrados en esta línea de investigación, en el último año se han finalizado 1 tesis de doctorado y 2 tesinas de grado de Licenciatura.

Actualmente se están desarrollando 4 tesis de doctorado, 1 tesis de especialista y 3 tesinas de grado de Licenciatura. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

- [1] Lanzarini L., Villa Monte A., Aquino G., De Giusti A. *Obtaining classification rules using lvqPSO*. Advances in Swarm and Computational Intelligence. Lecture Notes in Computer Science. Vol 6433, pp. 183-193. ISSN 0302-9743. Springer-Verlag Berlin Heidelberg. Junio 2015.
- [2] Jimbo P., Villa Monte A., Rucci E., Lanzarini L., Fernández A. *An exploratory analysis of methods for extracting credit risk rules*. XIII Workshop Bases de Datos y Minería de Datos (WBDDM). XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016), ISBN: 978-987-733-072-4, págs. 834-841, octubre de 2016
- [3] Jimbo Santana P., Villa Monte A., Rucci E., Lanzarini L., and Fernández Bariviera A.. *Analysis of Methods for Generating Classification Rules Applicable to Credit Risk*. Journal of computer science & technology (ISSN 1666-6038), vol. 17, num. 1, págs. 20-28, abril de 2017.
- [4] Lanzarini L., Villa Monte A., Fernandez Bariviera A., Jimbo Santana P. *Simplifying Credit Scoring Rules using LVQ+PSO*. : The International Journal of Systems& Cybernetics. Emerald Group Publishing Limited. vol. 46. Pp 8-16. ISSN 0368-492X. 2017.
- [5] Basgall, M. J., Hasperué, W., Estrebou C., Naiouf M. *Clustering de un flujo de datos usando MapReduce*. XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016). Pp 682-691. ISBN 978-987-733-072-4. Octubre 2016.
- [6] Basgall, M. J., Hasperué, W., Estrebou C., Naiouf M. *Data stream treatment using sliding windows with MapReduce*. Journal of Computer Science & Technology. Vol. 16. ISSN 1666-6038. Pp. 76-83. 2016.
- [7] Basgall, M. J., Hasperué, W., Naiouf, M., & Bariviera, A. F. (2017). *Cálculo del exponente de Hurst utilizando Spark Streaming: enfoque experimental sobre un flujo de transacciones de criptomonedas*. XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017).
- [8] Bariviera, A. F., Basgall, M. J., Hasperué, W., & Naiouf, M. (2017). *Some stylized facts of the Bitcoin market*. Physica A: Statistical Mechanics and its Applications, 484, 82–90. <https://doi.org/10.1016/j.physa.2017.04.159>
- [9] Villa Monte A., Lanzarini L., Rojas L., Oliva Varela J.. *Document Summarization using a Scoring-Based Representation*. 2016 XLII Latin American Computing Conference (CLEI), págs. 1-7, doi. 10.1109/CLEI.2016.7833396, octubre de 2016.
- [10] Villa Monte A., Lanzarini L., Rojas Flores L., Olivas Varela J. A.: *Document summarization using a scoring-based representation*. XLII Conferencia Latinoamericana en Informática (CLEI 2016). ISBN 978-1-5090-1633-4, pp. 1-7. Octubre de 2016.
- [11] Villa-Monte A., Lanzarini L., Fernández-Bariviera A. and Olivas J. A.. *Obtaining and evaluation of extractive summaries from stored text documents*. III Conference on

- Business Analytics in Finance and Industry.
Enero 2018. *En prensa*.
- [12] Puente C., Villa Monte A., Lanzarini L.,
Sobrino A. and Olivas Varela J. Á.
*Evaluation of causal sentences in automated
summaries*. Proceedings of the 2017 IEEE
International Conference on Fuzzy Systems
(FUZZ-IEEE), págs. 1-6, doi.
10.1109/FUZZ-IEEE.2017.8015666, 2017.
- [13] Coria, J.M. Clasificación de Subjetividad
utilizando Técnicas de Aprendizaje
Automático. Tesis de grado. Facultad de
Informática, UNLP. Febrero 2018.

Computación Gráfica, Imágenes y Visualización

Líneas de Investigación del Grupo de Realidad Aumentada Aplicada: Plantillas de Catálogos Aumentados Integración Escalable de Realidad Aumentada basada en Imágenes y Rostros Aumentación de Sistemas SCADA en el Contexto de la Industria 4.0

Martin Becerra¹, Nahuel Mangiarua¹, Santiago Igarza¹, Jorge Ierache¹, María José Abasolo^{2,3}

¹Universidad Nacional de La Matanza, DIIT, Grupo de Realidad Aumentada Aplicada
Florencio Varela 1903, La Matanza, Buenos Aires, Argentina

jierache@unlam.edu.ar

²Universidad Nacional de la Plata, Facultad de Informática, III-LIDI

³Comisión de Investigaciones Científicas de la Pcia. de Bs. As.

Resumen

Las líneas de investigación y desarrollo presentadas tienen por objetivo desarrollar aplicaciones de Realidad Aumentada en dirección a tres ejes. En primer lugar, se apunta al desarrollo de plantillas para la creación de catálogos aumentados que permitan la visualización de información añadida a partir de la captura de marcadores. En segundo lugar, se apunta a aplicaciones basadas en la detección de imágenes y rostros, y su integración y escalabilidad en el sistema de catálogos aumentados. Por último, se apunta empleo de Realidad Aumentada en el contexto de industrias 4.0, aumentando las capacidades de los sistemas SCADA de Supervisión, Control y Adquisición de Datos.

Palabras clave: Realidad Aumentada, Catálogo Virtual Aumentado, SCADA, Detección de Rostros, Industrias 4.0

Contexto

La investigación presentada es desarrollada por el grupo de investigación de Realidad Aumentada Aplicada del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza, en el marco del proyecto PROINCE C-202 2017-2018 PROINCE C-202 Framework para la Generación de Plantillas en Sistemas de

Catálogos de Realidad Aumentada. El grupo se encuentra financiado a través de las becas otorgadas por la institución, en particular el desarrollo de plantillas de catálogos aumentados por el proyecto antedicho.

Introducción

La Realidad Aumentada (RA) permite la fusión de datos virtuales sobre el mundo físico, enriqueciendo la percepción de la realidad, es decir aumentándola [1]. La Realidad Aumentada basada en el reconocimiento de imágenes planas consiste en utilizar como disparador un anclaje al mundo físico en forma de una imagen plana arbitraria que se detecta para la incorporación de información virtual de forma espacial y contextualmente coherente con la realidad. El marcador típico de RA consiste en una imagen cuadrada con una figura o patrón codificado de color blanco sobre fondo negro. Recientemente las aplicaciones de RA permiten utilizar como disparador una imagen arbitraria.

En los últimos años, la RA se ha expandido a diferentes campos de aplicación tales como educación, salud, turismo, marketing y entretenimiento. Este equipo de investigación ha desarrollado diversas aplicaciones, como juegos de tablero [2][3], herramientas para la generación de materiales didácticos para el área educativa [4][5], o sistemas de aumentación de información de salud mediante una tarjeta

aumentada basadas en conocimiento para la asistencia médica en emergencias [6][7]. El proyecto de investigación aplicada se encuadra en el contexto de la aplicación de tecnologías de RA en la vida cotidiana de las personas, generando así un impacto en la sociedad que contribuye a la participación de los mismos en el ámbito tecnológico.

Sistema de Catálogos Aumentados

Se desarrolló el Sistema de Catálogos Virtuales Aumentados [8] [9], el cual permite la generación, distribución y explotación de Catálogos de Realidad Aumentada por parte de usuarios finales. Dichos catálogos se componen por un conjunto de marcadores que son aumentados con información provista por los usuarios al momento de su creación, la cual es visualizada utilizando una aplicación para teléfono inteligente conectada a internet. El sistema de catálogos virtuales permite predefinir la cantidad y tipos de contenidos asociados a cada marcador junto con sus transformaciones geométricas (posición, rotación, escala) y su orden de aparición en el editor. Se pretende simplificar de la labor del usuario al momento de construir y generar contenido aumentado sin la necesidad de tener que contar con conocimientos específicos del dominio de la RA mediante el uso de templates.

Realidad Aumentada basada en Imágenes y Rostros

El sistema de catálogos aumentados desarrollado posibilita la creación de catálogos de tamaños manejables, permitiendo la utilización de marcadores pero no escapa de la limitación de escalabilidad cuando se utiliza RA basada en el reconocimiento de imágenes. En esta investigación se persigue incrementar la escalabilidad - es decir la cantidad de imágenes que pueden ser identificadas en una misma aplicación sin dependencia de un servicio web externo- de aplicaciones de RA basadas en reconocimiento de imágenes.

La búsqueda o identificación de imágenes particulares de entre un gran volumen de datos es un área de investigación muy activa y de rápida aplicación a nivel productivo como podemos observar con el florecimiento de servicios de búsqueda online inversa incluidos en buscadores web como Google. Es decir servicios que, a partir de una imagen dada como criterio de búsqueda, nos proporcionan otras similares o sitios web relacionados. En este contexto, los enfoques actuales consisten en el desarrollo de métodos de búsqueda aproximada predecibles, o en la reducción de dimensionalidad de la entrada o query. Por una parte, los primeros permiten el uso de entradas de gran tamaño al sacrificar de manera predecible la precisión de la búsqueda. Por otra parte, los segundos buscan reducir el tamaño de la entrada con una cierta pérdida de información, pero permitiendo el cálculo de funciones de distancia entre elementos más robusta y por ende costosa. Uno de los temas centrales de nuestra investigación es buscar una integración de técnicas de búsqueda a escala web existentes que cumpliendo con las restricciones de espacio y tiempo necesarios en el contexto de la RA, permitan incrementar la escalabilidad, comparada con sistemas actuales en términos de cantidad de imágenes distintas que pueden ser identificadas automáticamente sin el agregado de un marcador envolvente.

En particular, se propone trabajar sobre el indexado inteligente de descriptores locales tanto para imágenes como para registros biométricos, comenzando con técnicas simples como el indexado inverso jerárquico utilizando k-medios sobre los datos sin pre-procesar y continuar explorando en las técnicas más avanzadas citadas anteriormente.

Se plantea además la incorporación de detección de rostros como disparadores naturales para contenidos de RA, en particular sobre las bases del sistema de catálogo virtuales aumentados. Para ello, se trabaja en la integración de detectores públicamente

disponibles como el ya clásico detector en cascada propuesto por Violas P. y Jones M.[10]. Específicamente se busca la integración del entorno y motores de RA actualmente utilizados en el sistema de catálogos aumentados con la biblioteca OpenCV en un ambiente Android.

A su vez, esta línea de investigación pretende ser el puntapié inicial para futuros desarrollos que exploten la interacción de aumentación mediante el aprovechamiento de la información brindada por las imágenes de rostros humanos. La detección de rostros abre las puertas a un nuevo campo de aumentación de información tanto para usuario finales como para expertos. La incorporación de esta funcionalidad sentará las bases para el aprovechamiento de la información implícita en los rostros, desde contextos de identificación personal hasta afectivos. La inferencia de parámetros biométricos a partir de imágenes es un campo con numerosas especializaciones. Particularizando en la detección de emociones a través del rostro, encontramos un gran número de técnicas tanto a nivel experimental como a nivel productivo. Corneanu et al. [11] nos ofrece una extensiva introducción y recorrido por los más importantes aportes hasta la actualidad. Se pretende explorar las posibilidades de obtención e inferencia de información biométrica a través de la misma secuencia de video para su explotación en el contexto de la Realidad Aumentada. Dependiendo del ámbito de explotación, el estado emocional y otros factores como por ejemplo, el tamaño corporal, la posición de los miembros, etc., otorgan un nivel de interacción adicional de gran utilidad en futuras aplicaciones, en particular el contexto de computación afectiva.

Realidad Aumentada el Contexto de la Industria 4.0

Hasta la fecha la realidad aumentada ha sido reconocida en la industria como un apoyo

interesante para aplicaciones de mantenimiento, montaje y reparación de maquinaria [12]. Pero no existe hasta la fecha sistemas de RA que apoyen a humanos en la toma de decisiones en la supervisión de situaciones emergentes que podrían suceder en una planta industrial. El empleo de RA en el contexto de industrias 4.0 [13] permite aumentar las capacidades de los sistemas de monitoreo y control industriales. El objetivo de la línea de investigación de RA aplicada a sistemas en tiempo real en el contexto de la industria 4.0 se orienta al desarrollo de un prototipo que se integre a un sistema SCADA (Supervisory Control And Data Acquisition, en español Supervisión, Control y Adquisición de Datos)[14]. Se busca que el prototipo disponga de un sistema multiagente [15] para percibir eventos disparados por máquinas de una planta industrial, con el objetivo de diagnosticar situaciones o estados de las máquinas y que asista a los supervisores de planta en la toma de decisiones mediante el apoyo de sistemas basados en conocimiento. Se pretende que el prototipo proporcione una interfaz de RA para visualizar la síntesis de información observada y procesada por el sistema multiagente para realizar la supervisión descentralizada de equipos de la planta industrial. En una primer instancia se implementará para teléfonos móviles para luego poder migrar el visor a un sistema de gafas de retroproyección.

Líneas de Investigación, Desarrollo e Innovación

Las líneas de investigación y desarrollo presentadas tienen por objetivo desarrollar aplicaciones de Realidad Aumentada en dirección a:

- Templates de catálogos aumentados
- RA basada en reconocimiento de imágenes y rostros
- RA en el contexto de la industria 4.0

Resultados y Objetivos

En relación al desarrollo de catálogos aumentados se finalizó el módulo para la creación de templates para la plataforma web del Sistema de Catálogos Virtuales Aumentados. Este módulo le permitirá al usuario crear un nuevo template definiendo un nombre y una descripción al mismo, agregando marcadores con un nombre propio y para cada uno de ellos, asociar distintos tipos de contenidos con su orden, nombre propio y transformación geométrica. Finalmente se encuentra entre nuestras líneas trabajo la finalización de la implementación de grupos de usuarios y su integración a redes sociales con el objetivo final de lograr una beta abierta del sistema. Como objetivo se plantea la posibilidad de utilizar rostros humanos como marcadores de RA. Finalmente, se tiene como objetivo adicional la adaptación e integración del sistema de catálogos virtuales aumentados y los demás sistemas y aplicaciones del grupo para su correcto funcionamiento en dispositivos de vanguardia que potencian su utilidad y nos acercan al ambiente ideal para la explotación de la RA.

En relación a la línea de RA aplicada a SACADA se encuentra en los primeros pasos. Se pretende inicialmente la integración con un sistema de gafas de retroproyección en cristal transparente Moverio-300, del fabricante Epson. Así también hemos incorporado nuevos

dispositivos portátiles y visores, algunos con capacidad de RA y otros de inmersión en Realidad Mixta orientados al adiestramiento en infraestructuras industriales y sistemas de geoposicionamiento.

Formación de Recursos Humanos

El grupo de investigación se encuentra conformado por un investigador formado, tres investigadores en formación, y cuatro alumnos becarios del departamento de Ingeniería e Investigaciones Tecnológicas de la UNLaM, trabajando en el área de RA. Dos investigadores se encuentran realizando el Doctorado en Ciencias Informáticas en la UNLP, entre los cuales se enumera:

- Jorge Ierache, Nahuel Mangiarua, Martín Becerra, Santiago Igraza (UNLaM) . “Templates de Catálogos Aumentados”
- Nahuel A. Mangiarua “Integración escalable de Realidad Aumentada basada en imágenes y rostros” Directores: Jorge S. Ierache (UNLaM), María José Abásolo (UNLP)
- Martín Becerra “Aumentación de Sistemas SCADA en el Contexto de la Industria 4.0” Director: Jorge Ierache (UNLaM) en proceso de presentación a UNLP

Referencias

[1] Yee C., Abásolo M. J., Más Sansó R. y Vénere M. (2011). ”Realidad virtual y realidad aumentada. Interfaces avanzadas.” ISBN 978-950-34-0765-3.

[2] Ierache J., Mangiarua N., Verdicchio N., Sanz D., Montalvo C., Petrolo F., Igarza S. (2015). “Sistema de Catálogo Virtual Aumentado. Integración de Framework Especializado orientado a juegos didácticos” TEYET, pp 350-356, ISBN 978-950-656-154-

3.

[3] Verdicchio N., Sanz D., Igarza S., Mangiarua N., Montalvo C., Ierache J. (2016) "Sistema de Catálogo Virtual Aumentado Integración de Framework Especializado Orientado a Juegos Didácticos". TE&ET, pp 597-604, ISBN 978-987-3977-30-5.

[4] Mangiarua N., Ierache J., Bevacqua S., Becerra M., Verdicchio N., Duarte N., Sanz D., Igarza S. (2014), "Herramienta de Realidad Aumentada para la explotación de material didáctico tradicional". TE&ET. E-Book. ISBN 978-987-24611-1.

[5] Ierache J., Mangiarua N., Bevacqua S., Becerra M., Verdicchio N., Duarte N., Sanz D., Igarza S. (2014). "Herramienta de Realidad Aumentada para facilitar la enseñanza en contextos educativos mediante el uso de las TICs". Revista Latinoamericana de Ing de Software, 1(1): -3, ISSN 2314-2642.

[6] Ierache N., Mangiarua N., Verdicchio D., Sanz D., Montalvo C., Petrolo F. and Igarza S., "Augmented. Card System Based on Knowledge for Medical Emergency Assistance". IEEE Xplore ISBN 978-1-5090-2938-9 2016.

[7] Ierache J., Verdicchio N., Duarte N., Montalvo C., Petrolo F., Sanz D., Mangiarua N., Igarza S., "Augmented Reality Card System for Emergency Medical Services", IWBBIO 2016 Proceedings Extended abstracts 20 - 22 abril 2016 Granada (SPAIN), pp 487-494, ISBN 978-84-16478-75-0.

[8] Ierache J., Mangiarua N., Verdicchio N., Becerra M., Duarte N., Igarza S. (2014). "Sistema de Catálogo para la Asistencia a la Creación, Publicación, Gestión y Explotación de Contenidos Multimedia y Aplicaciones de Realidad Aumentada". CACIC 2014 Red UNCI ISBN 978-987-3806-05.

[9] Ierache J., Mangiarua N., Bevacqua S., Verdicchio N., Becerra M., Sanz D., Sena M., Ortiz F., Duarte N., Igarza S. (2015). "Development of a Catalogs System for Augmented Reality Applications". World Academy of Science, Engineering and Technology, International Science Index 97, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 9(1), 1 - 7. ISSN 1307:6892.

[10] Viola P., Jones M., Rapid Object Detection using a Boosted Cascade of Simple Features, 2001. [En línea]. Disponible en <https://goo.gl/GcGWaf>. [Accedido: 5 Marzo 2018].

[11] Corneanu, C. A., M. O. Simón, J. F. Cohn, and S. E. Guerrero. 2016. "Survey on RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition: History, Trends, and Affect-Related Applications." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38 (8):1548–68. <https://doi.org/10.1109/TPAMI.2016.2515606>.

[12] Azuma R. T., The Most Important Challenge Facing Augmented Reality, *Presence Teleoperators Virtual Environ.*, vol. 25, n.º 3, pp. 234-238, dic. 2016.

[13] Industria 4.0: La cuarta revolución industrial - guía a la Industria 4.0. [En línea]. Disponible en: <https://www.i-scoop.eu/industry-4-0/>. [Accedido: 5 marzo 2018].

[14] Daneels A., Salter W., «What is SCADA?», 1999.

[15] Russell S., Norvig P. , Inteligencia Artificial, un enfoque moderno, tercera edición 2010, cap. 2.

Métricas, Técnicas y Semántica para la Visualización de Datos

Silvia Castro^{1,2,3}, Martín Larrea^{1,2,3}, Dana Urribarri^{1,2}, Luján Ganuza^{1,2}, Sebastián Escarza^{1,2}

¹Laboratorio de I+D en Visualización y Computación Gráfica (UNS-CIC Prov. de Buenos Aires)

²Departamento de Ciencias e Ingeniería en Ciencias de la Computación (DCIC)

³Instituto de Ciencias e Ingeniería de la Computación (ICIC),

Universidad Nacional del Sur, Av. Alem 1253, Bahía Blanca

{smc, mll, dku, mlg, se}@cs.uns.edu.ar

RESUMEN

Un desafío importante al iniciar un proceso de visualización es el de encontrar una representación adecuada para un conjunto de datos y brindar las interacciones apropiadas. Esta línea de investigación tiene como objetivo estudiar nuevas alternativas para asegurar que el resultado de un proceso de visualización sea satisfactorio, tanto para aquel que crea la visualización como para el que la consume. Esta línea de investigación incluye el desarrollo y perfeccionamiento de un modelo unificado de visualización, métricas para la evaluación de técnicas de visualización, técnicas para dominios específicos de aplicación y la asistencia al usuario durante la creación de una visualización mediante la incorporación de información semántica y razonadores semánticos.

Palabras clave: *Ontología, Representación Formal, Visualización basada en semántica, Visualización.*

CONTEXTO

Este trabajo se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, <http://vyglab.cs.uns.edu.ar>) del Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur. Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales,

alumnos de grado y investigadores extranjeros. La línea de Investigación presentada está inserta en el proyecto acreditado *Análisis Visual de Grandes Conjuntos de Datos (24/N037)*, dirigido por la Dra. Silvia Castro y en el proyecto *Visualización Basada en Semántica (24/ZN29)* dirigido por el Dr. Martín Larrea; ambos financiados por la Secretaría General de Ciencia y Tecnología de la Universidad Nacional del Sur.

1. INTRODUCCIÓN

El objetivo de una visualización es encontrar una metáfora visual que permita entender y percibir en forma efectiva un conjunto de datos. Una visualización debe proveer también un conjunto de interacciones a partir de las cuales el usuario explorará el conjunto de datos con una mínima carga cognitiva. La tecnología computacional actual permite la exploración de grandes conjuntos de información; dichos conjuntos se hacen disponibles cada día a través de diferentes sistemas de información; estos datos varían entre datos no estructurados, documentos multimediales e información estructurada en base de datos. Por un lado, esta situación es extremadamente útil y excitante, pero la creciente cantidad de información genera una sobrecarga cognitiva llegando a un punto de ansiedad. Actualmente existe un gran número de

modelos de referencia en Visualización ([1, 2, 3]) que, de diversas maneras, han identificado y delineado los principales procesos y transformaciones que sufren los datos para ser visualizados. En particular, en el seno del VyGLab se ha desarrollado el Modelo Unificado de Visualización (MUV) [4], un modelo que constituye un marco conceptual de referencia en términos del cual ubicar los procesos y estados intermedios de los datos y definir las interacciones explícitamente. Si bien estos esfuerzos por compilar, caracterizar y clasificar los aspectos más relevantes del área han sido y siguen siendo fructíferos, aún no existe un consenso definitivo que permita consolidar una teoría de base en Visualización. Sobre este modelo hemos trabajado sobre la incorporación de semántica. La incorporación de una semántica bien definida en el proceso de Visualización ya ha sido identificada como una necesidad ([5, 6]) en el sentido que permitiría formalizar dicho proceso. A partir de esta formalización, se definirá un vocabulario común que permitirá a los usuarios plantear sus requerimientos, y a los diseñadores de herramientas de visualización, expresar las transformaciones de los datos. También sobre el MUV hemos trabajado en la definición de las interacciones sobre el proceso, tanto de alto como bajo nivel. Hemos incorporado al proceso de visualización técnicas específicas a dominios de aplicación, como es el caso de las Ciencias Geológicas y Captura de Movimiento. También, ante la necesidad de contar con métodos que nos permitan evaluar la calidad de una visualización, hemos desarrollado métricas para tal sentido.

2. LÍNEAS DE INVESTIGACIÓN Y

DESARROLLO

Este trabajo presenta tres ejes de investigación las cuales se entrelazan y tienen varios puntos en común. Tales ejes son:

1. “Visualización de Ciencias Geológicas”.
2. “Visualización Basada en Semántica”.
3. “Predicción del desempeño de las técnicas de visualización a partir de métricas sobre los datos”

Visualización de Ciencias Geológicas.

Este trabajo se enfoca en la visualización de un conjunto de datos geológicos, en particular, el conjunto de minerales que integran el grupo de los Espinelos. Este grupo de minerales resulta un excelente candidato a ser explorado y visualizado ya que es considerado un importante indicador petrogenético. Estos minerales proveen información vital en lo referido al ambiente tectónico de las rocas presentes en determinada área en el contexto de la tectónica global. En 2001, Barnes y Roeder ([7]) compilaron en una base de datos más de 26.000 análisis de Espinelos correspondientes a rocas ígneas y metamórficas. En base a estos análisis delinearon contornos para un conjunto de campos composicionales característicos (o patrones). Los geólogos suelen utilizar estos contornos para estimar el ambiente tectónico donde una muestra de Espinelos se podría haber formado. Esta tarea es propensa a errores e involucra una tediosa comparación manual de diagramas superpuestos. El objetivo de este trabajo de investigación es desarrollar una herramienta de Visualización que no solo permita visualizar y explorar conjuntos de datos correspondientes a composiciones minerales pertenecientes al grupo de los Espinelos, sino que también ofrezca técnicas que permitan categorizar

semiautomáticamente e interactivamente un conjunto arbitrario de Espinelos en función de los contornos de Barnes y Roeder.

Visualización Basada en Semántica. Este eje tiene como objetivo crear un modelo de visualización que considere la semántica de los datos, del contexto y de las etapas del proceso de visualización para poder asistir al usuario en el seteo de los parámetros de la visualización para lograr un resultado efectivo. Este nuevo modelo transformará datos en información, de acuerdo a Keller & Tergan [8], *“information is data that has been given meaning through interpretation by way of relational connection and pragmatic context”*. Los mismos datos pueden dar origen a diferente información, de acuerdo a la interpretación que se les de. Considerando esto, el nuevo proceso de visualización será capaz de determinar las características de una representación efectiva y podrá guiar al usuario a través de los diferentes estados del modelo. De esta manera, se logrará desarrollar un meta modelo para el proceso de visualización, los estados, los datos, las interacciones y el contexto.

Predicción del desempeño de las técnicas de visualización a partir de métricas sobre los datos. El objetivo de una visualización es obtener una representación del conjunto de datos que ayude al usuario en la correcta interpretación de los mismos y así lograr un acertado análisis de estos. Dado el constante crecimiento de los conjuntos de datos en diferentes y variados campos de la información, la tarea de elegir la técnica más adecuada para visualizar convenientemente los datos no es sencilla.

Además, el resultado del proceso de visualización depende de todas las decisiones que se hayan tomado a lo largo de dicho proceso: un usuario inexperto es propenso a tomar decisiones equivocadas afectando negativamente la visualización obtenida y, a la larga, frustrando su experiencia con la visualización. Si bien a la hora de visualizar conjuntos de datos pequeños no hay grandes desafíos, la situación cambia al intentar visualizar grandes conjuntos de datos: una mala decisión tomada en cualquier punto del proceso de visualización y el resultado obtenido puede no ser satisfactorio. Una alternativa para solucionar este problema es guiar al usuario en la toma de decisiones a lo largo del proceso. Sin embargo, esta tarea no es sencilla: implica la existencia de herramientas que permitan predecir qué decisión es “más conveniente” tomar. Una forma de elegir la decisión más conveniente es basarse en métricas sobre los datos que describen aspectos claves de la técnica y permitan predecir el resultado final sin necesidad de aplicar la técnica sobre los datos.

3. RESULTADOS Y OBJETIVOS

Publicaciones realizadas bajo el eje de investigación “**Visualización de Ciencias Geológicas**”: [9, 10, 11, 12, 13, 14]

Publicaciones realizadas bajo el eje de investigación “**Visualización Basada en Semántica**”: [15, 16, 17, 18, 19, 20, 21].

Publicaciones realizadas bajo el eje de investigación “**Predicción del desempeño de las técnicas de visualización a partir de métricas sobre los datos**”: [22, 23]

4. FORMACIÓN DE RECURSOS

HUMANOS

En lo concerniente a la formación de recursos humanos se detallan las tesis en desarrollo y concluidas relaciones con las líneas de investigación presentadas, así también como un proyecto en ejecución:

Tesis en Desarrollo

Tesis Doctoral. Sebastián Escarza. Tema: Ontologías de Visualización. Dirección: Dra. Silvia Castro.

Tesis Magister. Georgina Inés Cerúsico. Tema: Taxonomías de Técnicas de Visualización para la Formalización del Proceso de Visualización. Dirección: Dra. Silvia Castro y Dr. Martín Larrea.

Tesis Magister. Alejandra Elizabeth Herrera. Tema: Coordenadas Paralelas. Visualización e Interacciones. Dirección: Dra. Silvia Castro.

Tesis Magister. César Escobal Blanco. Tema: La Reducción de Datos en las Técnicas de Visualización. Dirección: Dra. Silvia Castro.

Tesis Concluidas

Tesis Doctoral. María Luján Ganuza. Tema: Servicios Web en Visualización. Dirección: Dra. Silvia Castro.

5. BIBLIOGRAFÍA

[1] E. H. Chi. A taxonomy of visualization techniques using the data state reference model. In Proceedings of the IEEE Symposium on Information Visualization (InfoVis'00), page 6975. IEEE Computer Society Press, 2000.

[2] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. IEEE Symposium on Visual Languages, 0:336, 1996.

[3] K. Brodlie and N. M. Noor. Visualization Notations, Models and Taxonomies. pages 207–212, Bangor, United Kingdom, 2007. Eurographics Association.

[4] Martig S., Castro S., Fillotrani P., Estévez E., Un Modelo Unificado de Visualización, Proceedings, pp. 881-892, 9° Congreso Argentino de Ciencias de la Computación. 6 al 10 de Octubre de 2003. La Plata. Argentina.

[5] D. J. Duke, K. W. Brodlie, D. A. Duce, and I. Herman. Do you see what I mean? IEEE. Computer Graphics and Applications, 25(3):6–9, 2005

[6] M. Chen, D. Ebert, H. Hagen, R. S. Laramee, R. van Liere, K.-L. Ma, W. Ribarsky, G. Scheuermann, and D. Silver. Data, information, and knowledge in visualization. IEEE Comput. Graph. Appl., 29(1):12–19, 2009.

[7] Barnes, S. J.; Roeder, P.L. 2001. The Range of Spinel Compositions in Terrestrial Mafic and Ultramafic Rocks. Journal of Petrology, vol. 42, number 12, pp: 2279-230

[8] Tergan, S. O., & Keller, T. (Eds.). (2005). *Knowledge and information visualization: Searching for synergies* (Vol. 3426). Springer.

[9] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E., Gröller, E., & Matković, K. (2014). The spinel explorer—interactive visual analysis of spinel group minerals. *IEEE transactions on visualization and computer graphics*, 20(12), 1913-1922..

[10] Ganuza, M. L., Gargiulo, F., Ferracutti, G., Castro, S., Bjerg, E., Gröller, E., & Matković, K. (2015, October). Interactive semi-automatic categorization for spinel group minerals. In *Visual Analytics Science and Technology (VAST), 2015 IEEE Conference on* (pp. 197-198). IEEE.

[11] Ferracutti, G. R., Gargiulo, M. F., Ganuza, M. L., Bjerg, E. A., & Castro, S. M. (2015). Determination of the spinel group end-members based on electron microprobe analyses. *Mineralogy and Petrology*, 109(2),

153-160.

[12] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E., Gröller, E., & Matković, K. (2014). The spinel explorer—interactive visual analysis of spinel group minerals. *IEEE transactions on visualization and computer graphics*, 20(12), 1913-1922.

[13] Immersive Analytics for Geology: Field Sketch-Like Visualization to assist geological structure analysis during fieldwork. IEEE VIS - Workshop on Immersive Analytics. N. F. Gazcón, J. M. Trippel, J. I. Larregui, M. L. Ganuza, E. A. Bjerg y S. M. Castro. Phoenix, USA.

[14] Interactive Visual Categorization of Spinel-Group Minerals. 33rd Spring Conference on Computer Graphics (SCCG 2017). M. L. Ganuza et al. Mikulov, República Checa.

[15] S. Escarza, S. Castro, S. Martig. Formalización del Proceso de Visualización mediante el uso de Ontologías. XII Workshop de Investigadores en Ciencias de la Computación (WICC 2010) pp. 261-265. Universidad Nacional de la Patagonia San Juan Bosco. 5 y 6 de mayo de 2010. El Calafate. Santa Cruz. Argentina. ISBN: 978-950-34-0652-6.

[16] M. Larrea, S. Martig and S. Castro. Semantics-based Color Assignment in Visualization. *Journal of Computer Science & Technology*. Vol. 10 - No. 1 – April 2010 - ISSN 1666-6038.

[17] M. Larrea, S. Martig and S. Castro. Formalización del Proceso de Visualización Basada en Semántica. XII Workshop de Investigadores en Ciencias de la Computación (WICC 2010) pp. 270-274. Universidad Nacional de la Patagonia San Juan Bosco. 5 y 6 de mayo de 2010. El Calafate. Santa Cruz. Argentina. ISBN: 978-950-34-0652-6.

[18] Escarza, S., Larrea, M. L., Urribarri, D. K., Castro, S. M., & Martig, S. R. (2011).

Integrating semantics into the visualization process. In *Dagstuhl Follow-Ups* (Vol. 2). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

[19] Larrea, M. L., Escarza, S., Urribarri, D. K., Castro, S. M., Herrera, A. E., Escobal Blanco, C. A., & Cerúsico, G. (2014, October). Ontologías y Semántica en el Proceso de Visualización. In *XVI Workshop de Investigadores en Ciencias de la Computación*.

[20] Larrea, M. L. (2017). Black-Box Testing Technique for Information Visualization. Sequencing Constraints with Low-Level Interactions. *Journal of Computer Science & Technology*, 17.

[21] Larrea, M., Castro, S., Urribarri, D., & Escarza, S. Visualización Basada en Semántica. *WICC 2017*, 418.

[22] Urribarri, D. K. (2015, August). Predicción del desempeño de las técnicas de visualización a partir de métricas sobre los datos. In *XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015)*

[23] Urribarri, D. K., & Castro, S. M. (2017). Prediction of data visibility in two-dimensional scatterplots. *Information Visualization*, 16(2), 113-125.

Tecnologías Inmersivas Aplicadas: Realidad Virtual y Aumentada

Matías Selzer^{1,2}, Nicolás Gazcón^{1,3}, Juan Trippel Nagel^{1,3}, Martín Larrea¹, Silvia Castro¹ y Ernesto Bjerg⁴

¹ Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (UNSCIC Pcia. de Buenos Aires)

Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur

² Comisión de Investigaciones Científicas (CIC)

³ Instituto de Ciencias e Ingeniería de la Computación (CONICET-UNS)

⁴ INGEOSUR y Departamento de Geología, Universidad Nacional del Sur

{matias.selzer, nfg, juan.trippel, mll, smc}@cs.uns.edu.ar, ebjerg@ingeosur-conicet.gob.ar

Resumen

El continuo avance del hardware permite que tecnologías inmersivas se vuelvan accesibles para su aplicación en tareas cotidianas. Hoy en día es común que una persona posea un *smartphone* o una *tablet*, con características comparables al poder de una computadora de escritorio. De esta manera, tanto el campo de la Realidad Aumentada (RA) como el de la Realidad Virtual (RV) pueden proveer aplicaciones para tareas cotidianas.

En esta línea de investigación se está trabajando en distintas aplicaciones prácticas de estas tecnologías mediante el uso de dispositivos móviles. Puntualmente en el uso de RA para aplicaciones científicas, de turismo y de educación. En el caso de la RV, presentamos su uso para el estudio de interacciones en ambientes virtuales en aplicaciones para la educación.

Palabras clave: *Realidad Aumentada, Realidad Virtual, Dispositivos Móviles, Interacción Humano Computadora.*

Contexto

La línea de investigación presentada está inserta en el proyecto “Análisis Visual de Grandes Conjuntos de Datos” (24/N037), dirigido por la Doctora Silvia Castro. El proyecto es financiado por la Secretaría General de Ciencia y Tecnología de la Universidad Nacional del Sur; y acreditado

por la Universidad Nacional del Sur, Bahía Blanca.

Introducción

Gracias al avance de la tecnología, es posible que dispositivos móviles con los requerimientos de *hardware* necesarios para ser utilizados en aplicaciones de Realidad Aumentada (RA) y Realidad Virtual (RV), estén al alcance del público general. Esto permite que se explore y se investigue su utilización en diversos campos, teniendo en cuenta las necesidades cotidianas de estos.

La RA permite que elementos virtuales sean introducidos en nuestro entorno real [1]. Este beneficio se ha explorado en diferentes campos [2, 3, 4], continuando el desafío de buscar nuevas aplicaciones concretas de esta tecnología. Dentro de esta línea de investigación exploramos su utilización en actividades científicas desarrolladas en tres áreas. Estas comprenden el área de las Ciencias Geológicas, el uso de RA para el turismo regional y su utilización en la educación.

En cuanto a la RV, aunque su aplicación ha sido centrada inicialmente en el terreno del entrenamiento y de los videojuegos, se ha extendido a muchos otros campos, como la medicina [5, 6, 7], la arqueología [8, 9], el entrenamiento militar [10] o diferentes tipos de simulaciones [11]. Dentro de esta línea de investigación

exploramos su utilización para la educación en el área de Geografía y Turismo locales.

Líneas de Investigación y Desarrollo

En el contexto de esta línea de investigación se están desarrollando los siguientes trabajos en paralelo:

1. Visualización de Datos Geológicos mediante RA y dispositivos móviles en el trabajo de campo.
2. RA aplicada al turismo y educación regional.
3. Interacciones en RV y sus aplicaciones en la educación.

Visualización de Datos Geológicos mediante RA y dispositivos móviles en el trabajo de campo.

Este trabajo es llevado a cabo por becarios doctorales y postdoctorales del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) del Dpto. de Ciencias e. Ingeniería de la Computación, Universidad Nacional del Sur (UNS) con asistencia del INGEOSUR CCT-CONICET. Una de las actividades que realizan habitualmente los geólogos consiste en trabajos de campo. Estas tareas se efectúan en ambientes exteriores, y en muchos casos en entornos que pueden ser hostiles y en los que sólo se tiene comunicación a través de los satélites.

La RA aplicada a las ciencias geológicas busca asistir en la práctica de dichas tareas, simplificando y complementando el uso de herramientas propias de la actividad como lo son las cartas geológicas, las brújulas y los mapas. Para esto se incorpora a la visión de la realidad, información que pueda resultar de utilidad para el geólogo como lo es, por ejemplo, información de formaciones geológicas u objetos virtuales que podrían indicar puntos de interés del geólogo. Las interacciones que se pueden lograr con la

visualización de datos en tiempo real en el mismo trabajo de campo pueden contribuir efectivamente a esta disciplina, aportando soluciones y resultados en base al entendimiento de los datos y a sus relaciones.

Debido a la naturaleza móvil y en entornos abiertos, el subsistema de *tracking* es de suma importancia para asistir correctamente al geólogo [12]. En este sentido los sistemas de navegación (GPS, GLONASS) permiten la obtención de la ubicación del sistema en cualquier punto del globo con una precisión lo suficientemente buena para satisfacer los requerimientos de *tracking* de posición. Por otro lado, los sensores inerciales y magnéticos, provistos en la mayoría de los dispositivos móviles modernos (*Tablets* o *SmartPhones*), proveen el mecanismo de obtención de la orientación.

RA aplicada al turismo y educación.

Este trabajo se lleva a cabo con investigadores del VyGLab del Dpto. de Ciencias e. Ingeniería de la Computación, Universidad Nacional del Sur (UNS), en colaboración con investigadores del Departamento de Geografía y Turismo de la UNS.

Este trabajo se enfoca en la inclusión de las TICs como complemento motivacional y de aprendizaje para la concientización de la importancia de los Humedales Costeros de Bahía Blanca. Este espacio regional resulta de gran importancia para su comunidad y para la preservación del ecosistema que comprende. Esto abre una posibilidad muy interesante para el uso de tecnologías inmersivas como la RA para la difusión de su importancia.

Esta línea también se ubica en el contexto del Proyecto “Innovación educativa: Los Humedales Costeros de la Bahía Blanca como Espacio Recreativo”, de la Comisión de Investigaciones Científicas

de la Provincia de Buenos Aires (PIT-AP-BA).

Interacciones en RV y sus aplicaciones en la educación.

Este trabajo se lleva a cabo con investigadores del VyGLab del Dpto. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (UNS).

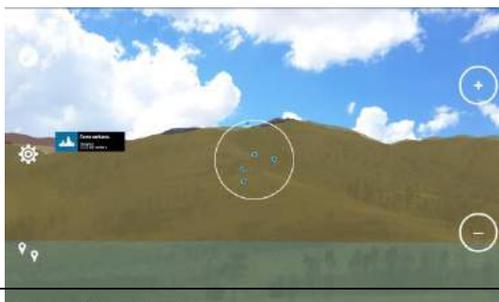
Actualmente las líneas de investigación se focalizan en los distintos tipos de interacciones que se pueden producir entre el usuario y el entorno virtual, considerando ambientes virtuales generados desde dispositivos de bajo costo. Las interacciones son evaluadas mediante la aplicación del modelo de Norman, bajo el desarrollo de casos de aplicación y orientadas a aplicaciones de educación.

Resultados y Objetivos

Visualización de Datos Geológicos mediante RA y dispositivos móviles en el trabajo de campo.

En esta línea se está comenzando con el desarrollo de un *framework* destinado a la visualización *in situ* de información geológica. Este soporta la inclusión de terrenos en 3D que puedan ser superpuestos a la vista del mundo real [2,3]. De esta manera se ofrecerán soluciones a los requerimientos básicos de un sistema de RA, esto es, los subsistemas de tracking y registración, visualización e interacción.

Actualmente se logró unificar la visualización de una superficie 3D, generada a partir de un mapa de altura, con la vista del terreno real (ver Figura 1). Para el subsistema de tracking se utilizaron la tecnología GPS y GLONASS para



26 y 27 de Abril de 2018
Figura 1. Visualización del terreno 3D y de un punto de interés.

determinar la posición y una fusión de sensores inerciales y magnéticos para la obtención de la orientación.

En cuanto a la visualización del terreno se desarrolló una librería basada en OpenGL para aprovechar las capacidades provistas por las GPU integradas en los dispositivos móviles.

El objetivo es utilizar este *framework* para desarrollar una aplicación que asista al geólogo en el trabajo de campo. Pudiendo almacenar puntos de interés geo-referenciados con la posibilidad de obtener información del contexto.

Actualmente las líneas de investigación se focalizan en los distintos tipos de interacciones que se pueden producir entre el usuario y el entorno virtual, considerando ambientes virtuales generados desde dispositivos de bajo costo. Las interacciones son evaluadas mediante la aplicación del modelo de Norman, bajo el desarrollo de casos de aplicación.

De esta manera se aportará una solución al problema de la ubicación relativa de los distintos puntos de interés respecto a la ubicación actual del usuario en el mundo.

Con estas nuevas funcionalidades el geólogo podría agilizar el trabajo de campo procesando datos in-situ, sin necesidad de realizarlo en un trabajo posterior de oficina.

RA aplicada al turismo y educación.

En esta línea se ha realizado un primer prototipo de una aplicación de RA que fue utilizada en Villa del Mar, localidad de Punta Alta, Provincia de Buenos Aires. La aplicación, denominada *HumedalesAR*, permite al usuario explorar la zona mostrando puntos de interés con información relevante a estos (texto, enlaces a la web, contenido multimedia y juegos didácticos). Al explorar la zona, la aplicación informa la distancia a los puntos de interés, y a medida que el usuario se

acerca a estos puede interactuar con la información que se provee en función de la distancia al punto de interés (ver Figura 2).

La aplicación fue utilizada dentro de un recorrido en la zona por cuatro cursos (cada curso compuesto por 25 alumnos aproximadamente). En esta prueba se pudieron obtener resultados para mejorar la aplicación, que será utilizada nuevamente en el corriente año.



Figura 2: *HumedalesAR*: (a) Ejemplo de dos puntos de interés y (b) ejemplo al estar posicionado en las inmediaciones de un punto de interés.

Interacciones en RV y sus aplicaciones en la educación.

Nuestra línea de investigación consiste en la investigación y el desarrollo de aplicaciones de RV enfocadas especialmente para funcionar en dispositivos móviles. Nos centramos principalmente en técnicas para reducir el costo computacional de los elementos gráficos y, en relación a la línea de investigación anterior, investigar sobre los distintos tipos de interacciones que se pueden lograr entre los usuarios y los dispositivos móviles, aprovechando al máximo las herramientas que nos proveen,

como por ejemplo sus bastos sensores y sus posibilidades de intercomunicación.

Actualmente se está desarrollando un prototipo para incorporar interacciones a un entorno virtual, que permita evaluar el impacto de la RV en el aprendizaje de conceptos. La Figura 3 muestra un ejemplo del prototipo en desarrollo, constando de una aplicación que permita que varios alumnos participen en un entorno virtual, al mismo tiempo que el docente puede seguir el avance de estos y disparar eventos o interacciones relevantes al entorno virtual utilizado.

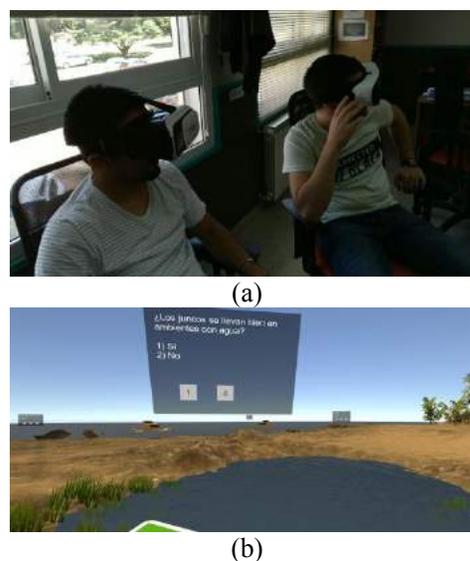


Figura 3: (a) Ejemplo de cascos de RV de bajo costo que pueden utilizarse mediante los teléfonos y (b) captura del prototipo en actual desarrollo del entorno virtual.

Formación de Recursos Humanos

En lo concerniente a la formación de recursos humanos se detallan las tesis concluidas y en desarrollo relacionadas con la línea de investigación presentada:

Tesis en Desarrollo

- Tesis Doctoral. Matías Selzer. Tema: *Métricas de Inmersión para Sistemas de Realidad Virtual*. Dirección: Dra. Silvia Castro, Dr. Martín Larrea.

- Tesis Doctoral. Juan Manuel Trippel Nagel. Tema: *Realidad Aumentada Móvil en Exteriores para Visualización de datos Geológicos*. Dirección: Dra. Silvia Castro, Dr. Ernesto Bjerg.

Referencias

- [1] Ronald Azuma, A survey of augmented reality, *Presence: Teleoperators & Virtual Environments* 6 (4), 355-385.
- [2] Fedorov R., Frajberg D., Fraternali P. A Framework for Outdoor Mobile Augmented Reality and Its Application to Mountain Peak Detection. *AVR* (1) 2016: 281-301.
- [3] L. Frauciel, J. Vairon, P. Nehlig, P.Thierry, I. Zendjebil, F. Ababsa. Outdoor Augmented Reality as a tool for bringing 3D geology to the field: the RAXENV project.
- [4] Koutek, M., 2003. Scientific Visualization in Virtual Reality: Interaction Techniques and Application Development. Computer Graphics & CAD/CAM group, Faculty of Information Technology and Systems (ITS), Delft University of Technology (TU Delft).
- [5] Freedman, S. A., Dayan, E., Kimelman, Y. B., Weissman, H., & Eitan, R. (2015). Early intervention for preventing posttraumatic stress disorder: an Internet-based virtual reality treatment. *European journal of psychotraumatology*, 6.
- [6] Rothbaum, B. O., Price, M., Jovanovic, T., Norrholm, S. D., Gerardi, M., Dunlop, B. & Ressler, K. J. (2014). A randomized, double-blind evaluation of D-cycloserine or alprazolam combined with virtual reality exposure therapy for posttraumatic stress disorder in Iraq and Afghanistan War veterans. *American Journal of Psychiatry*.
- [7] Gorini, A., & Riva, G. (2014). Virtual reality in anxiety disorders: the past and the future. *Expert Review of Neurotherapeutics*.
- [8] Lynch, J., & Corrado, G. (2014). *Arqueología virtual aplicada al sitio Villavil, Catamarca, Argentina* Virtual Archaeology applied to the site Villavil, Catamarca, Argentina.
- [9] Gaugne, R., Gouranton, V., Dumont, G., Chauffaut, A., & Arnaldi, B. (2014). Immersia, an open immersive infrastructure: doing archaeology in virtual reality. *Archeologia e Calcolatori*, supplemento 5, 1-10.
- [10] Carroll, J. M. (Ed.). (2003). *HCI models, theories, and frameworks: Toward a multidisciplinary science*. Morgan Kaufmann.
- [11] Schreuder, H. W., Persson, J. E., Wolswijk, R. G., Ihse, I., Schijven, M. P., & Verheijen, R. H. (2014). Validation of a novel virtual reality simulator for robotic surgery. *The Scientific World Journal*, 2014.
- [12] E. Veas, R. Grasset, E. Kruijff, D. Schmalstieg. *Extended Overview Techniques for Outdoor Augmented Reality*, *IEEE Trans. on Visualization and Computer Graphics* (Proceedings Virtual Reality 2012), Vol. 18, pp. 1-12, April 2012.

Plataforma Detección de objetos en tiempo real

Raúl Moralejo^{1,2}, Antonio Storni².

¹Escuela de Tecnología de la Información y las Comunicaciones
Departamento de Básicas y Tecnológicas
UNdeC - Universidad Nacional de Chilecito
romoralejo@gmail.com

²GridTICs – Grupo en Tecnologías de la Información y las Comunicaciones
Facultad Regional Mendoza
Universidad Tecnológica Nacional
romoralejo@gmail.com, antoniostorni@gmail.com

RESUMEN

En los últimos años, los sistemas de visión artificial han evolucionado sustancialmente. Hoy en día, la visión artificial permite identificar personas, lugares y objetos en imágenes con una precisión equivalente o superior a la capacidad humana y con mayor velocidad y eficiencia. Los métodos que utilizan Deep Learning han obtenido un performance superior a otros enfoques en el campo del reconocimiento de imágenes. El objetivo de esta investigación es implementar herramientas que resuelvan problemas complejos del mundo real aplicando las últimas técnicas de visión artificial, utilizando software libre.

Palabras clave:

Visión artificial, machine learning, deep learning.

CONTEXTO

El presente trabajo está inserto en el proyecto de investigación presentado en la Universidad Nacional de Chilecito, Departamento de Básicas y Tecnológicas. Escuela de Tecnología de la Información y las Comunicaciones, y en la Universidad Tecnológica Nacional, Facultad Regional Mendoza, Grupo GridTICs.

1. INTRODUCCION

Se puede definir la “Visión Artificial” como un campo de la “Inteligencia Artificial” que, mediante la utilización de técnicas adecuadas, permite la obtención, procesamiento y análisis de cualquier tipo de información espacial obtenida a través de imágenes digitales [1].

Las técnicas de visión artificial están presentes desde hace más de 40 años, pero no fué hasta principios de la década del 2000 que empezaron a aplicarse de modo masivo.

Los primeros algoritmos funcionan utilizando características codificando manualmente los objetos para hacer un reconocimiento, por ejemplo, se codificaban patrones para la nariz y los ojos para la detección de un rostro humano [2] [3].

Con el avance de las tecnologías de Machine Learning, y en especial Deep Learning, las herramientas de visión artificial han avanzado en su velocidad y precisión.

Actualmente, en vez de codificar manualmente las características de los objetos que queremos detectar, lo que necesitamos es

una gran cantidad de datos etiquetados para “entrenar” un clasificador basado en redes neuronales.

El gran avance en los últimos años en la precisión de las técnicas de Deep learning, se debe en parte debido al incremento en la capacidad de cómputo, y por otra parte, a la mayor disponibilidad de grandes sets de datos para entrenar las redes neuronales.

Por otra parte, las primeras aplicaciones de visión artificial, se hacía solamente una clasificación de imágenes, partiendo de una entrada de una imagen, podemos obtener un resultado textual de la clasificación de la imagen, por ejemplo, “perro”, “persona”, “bicicleta”.

Las nuevas técnicas de visión artificial no son sólo “clasificadores”, sino que también son “detectores”. Esto significa que además de la clase del objeto, podemos obtener su ubicación relativa a otros objetos y su tamaño, colocando en un cuadro delimitador [4][5].

Éstas técnicas de “detección” de objetos, fueron mejorando con los años, y actualmente podemos obtener detecciones de objetos sobre imágenes en un tiempo aproximado de 20 milisegundos, utilizando hardware de costo relativamente accesible. Esto implica que se puede hacer detección de video en tiempo real a una velocidad de 50 cuadros por segundo [6].

2. LINEAS DE INVESTIGACION y DESARROLLO

Los ejes principales son visión artificial, machine learning, deep learning, software de base libre (lenguaje de programación, framework, base de datos, sistema operativo) para el desarrollo de plataforma.

3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo principal del proyecto es generar una plataforma para la detección de objetos en tiempo real para distintos ámbitos.

Existen ya modelos entrenados y listos para usar, para reconocer objetos comunes, como, “persona, auto, bicicleta, camión”. En el caso de necesitar el reconocimiento de objetos específicos que no se encuentren en los modelos pre-entrenados, sólo se necesitaría contar con un set de datos etiquetados suficientemente grande y diverso, con imágenes de los objetos que queremos que el sistema reconozca para re-entrenar la red.

La plataforma podrá ser usada de la siguiente manera:

Configurando una o varias cámaras IP como entrada, definir o acotar un área para aplicar la detección de objetos en tiempo real, la clase de objeto que queremos detectar (por ejemplo: “persona”) y a través de un sistema de reglas condicionales, ejecutar distintas acciones que podrían ser: encender alarma, activar enviar un email o un mensaje instantáneo, aumentar o disminuir una variable de control, etc.

Beneficios de uso

Los beneficios de uso podrían dividirse en 3 categorías:

Medio Ambientales

El proyecto ayudará a disminuir las emisiones de CO2 al mejorar la fluidez del tránsito urbano.

Sociales

Puede ayudar a las personas con alguna discapacidad o en zonas en donde la inseguridad es un peligro latente.

Técnicos

El proyecto contribuirá a expandir el uso de técnicas de deep learning en la comunidad.

Casos de uso

Industria:

Definir una zona segura o insegura para que transiten las personas.

- Reconocer objetos permitidos y no permitidos en una cinta transportadora.

Seguridad:

- Definir zonas y horarios en los cuales si se detecta una persona, disparar una alarma.
- Detectar el rostro de personas autorizadas o alertar en caso de rostros no autorizados.

Ciudades:

- Sincronizar y optimizar los semáforos dependiendo de la cantidad de autos o peatones que haya en un momento dado.
- Generar datos estadísticos de movimiento de personas, vehículos, etc.

Transporte:

- Generación de un sistema para alertar o frenar un auto o tren en caso de que una persona se encuentre en el camino.

Tecnología de información utilizada

- Framework de Deep Learning desarrollado por Google, Tensorflow [7].
- Red neuronal convolucional pre-entrenada YOLO [8].
- Para comando de la red neuronal y procesamiento de imágenes, el lenguaje de programación Python [9].
- Para plataforma de administración web, el framework Django/Python [10].
- Base de datos PostgreSQL [11].
- Sistema operativo Linux, Ubuntu Server [12].

Principales pantallas del aplicativo

Carga de usuarios

Se pueden cargar usuarios y grupos, con permisos por usuario y por grupo.



Carga de cámaras



Ejemplos de detección

Ejemplo de Detección de personas y camión



Ejemplo de detección de automóvil y persona



Ejemplo de detección de persona y objeto (botella)



4. FORMACION DE RECURSOS HUMANOS

Uno de los principales objetivos del proyecto es la capacitación de los recursos humanos.

La meta como investigadores es la capacidad para realizar investigación científica, generar conocimiento y facilitar la transferencia de tecnología a la sociedad.

Este proyecto de investigación posibilita la colaboración inter-institucional y la ejecución entre grupos de I+D.

Para lograr estos objetivos se dispone del siguiente personal:

- 1 Investigador formado.
- 2 Investigadores de apoyo.

Adicionalmente se realizarán:

Dictado de cursos, seminarios y conferencias.
Promoción, coordinación y asistencia técnica de tesinas para alumnos de la carrera de grado de ingeniería y licenciatura en sistemas de la UNdeC.

Promoción, coordinación, dirección y asistencia técnica a tesis doctorales, postgrado y/o maestría.

Presentación de trabajos en congresos y reuniones científicas/técnicas.

Publicación de trabajos en revistas con/sin referato.

5. BIBLIOGRAFIA

- [1] VISIÓN ARTIFICIAL | Edgard Martinez
http://www.academia.edu/6084688/VISI%C3%93N_ARTIFICIAL
- [2] Viola Jones Algorithm.
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.10.6807>
- [3] HOG Algorithm.
<http://lear.inrialpes.fr/pubs/2006/DTS06/eccv2006.pdf>
- [4] FAST R-CNN (2015)
<https://arxiv.org/abs/1504.08083>

[5] YOLO (2015)
<https://arxiv.org/abs/1506.02640>

[6] YOLO Darknet implementation.
<https://pjreddie.com/darknet/yolo>

[7] <https://www.tensorflow.org/>

[8] <https://pjreddie.com/darknet/yolo/>

[9] <https://www.python.org/>

[10] <https://www.djangoproject.com/>

[11] <https://www.postgresql.org/>

[12] <https://www.ubuntu.com/>

Realidad Aumentada y Visión por Computador. Framework multipropósito

Alejandro Mitaritonna^{1,2}, Juan Lestani¹, Francisco Tarulla¹, Tomás Poeta¹, Silvana Olmedo¹,
Carolina Paez¹, Martín Lorenzo¹, Florencia Vela¹, Leonardo Inza¹

¹Laboratorio de Procesamiento de Imágenes y Visión Artificial
Universidad de Belgrano, Facultad de Ingeniería y Tecnología Informática
{alejandro.mitaritonna, juan.lestani, francisco.tarulla, tomas.poeta, silvana.olmedo, carolina.paez,
martin.lorenzo, florencia.vela, leonardo.inza}@comunidad.ub.edu.ar

²Instituto de Investigaciones Científicas y Técnicas para la Defensa (CITEDEF)
amitaritonna@citedef.gob.ar

RESUMEN

La línea de investigación consiste en el diseño y desarrollo de un framework de software multipropósito de Realidad Aumentada (RA) y Visión Artificial (VA) que se está llevando a cabo en el Laboratorio de Procesamiento de Imágenes y Visión Artificial de la Facultad de Ingeniería y Tecnología Informática de la Universidad de Belgrano (FITI-UB). Dicho framework servirá como base para la construcción de aplicaciones móviles y de escritorio para ser utilizadas en diferentes disciplinas. Uno de los principales objetivos del proyecto de investigación es la formación de recursos humanos y el fortalecimiento de las relaciones de I+D+i dentro y fuera de la FITI-UB ya sea con pares académicos y con la industria privada.

Palabras claves: *realidad aumentada, visión por computador, framework*

CONTEXTO

El diseño y desarrollo del framework surge como iniciativa de la FITI-UB con el fin de crear un laboratorio de Procesamiento de

Imágenes y Visión Artificial dedicado a la investigación y desarrollo de proyectos relacionados con la temática. Dentro de esta línea se está culminando una tesis doctoral en el Instituto de Investigaciones Científicas y Técnicas para la Defensa (CITEDEF) en el marco del Programa de Investigación y Desarrollo para la Defensa (PIDDEF) (aprobado con el Nro. 22/12), elaborado por la Subsecretaría de Investigación Científica y Desarrollo Tecnológico del Ministerio de Defensa.

INTRODUCCION

La línea de investigación presentada se inscribe dentro de las áreas de procesamiento de imágenes y visión por computador. En función de las necesidades de aprendizaje y de investigación que forman parte del diseño y desarrollo del framework propuesto se han creado 3 grupos de investigación conformados por alumnos de tercer año, cuarto y quinto año de las carreras de Licenciatura en Sistemas de Información e Ingeniería en Informática de la FITI-UB. Estas líneas/grupos de investigación se centraron en investigar y experimentar sobre:

estimación de posición y orientación (pose) y registración de objetos, reconocimiento y clasificación de objetos y motores de renderizado de objetos. En [Mit7] puede encontrarse una introducción del proyecto de investigación y, como parte de la experiencia adquirida en la construcción de frameworks de RA y VA, se puede encontrar un detalle en [Mit15].

Realidad Aumentada

Un sistema de RA es un tipo de sistema de visión sintética que mezcla gráficos generados por una computadora (anotaciones) con el mundo real. Las anotaciones proporcionan información destinada a ayudar a la toma de decisiones. El sistema de RA debe decidir qué anotaciones mostrar y cómo mostrarlas para asegurar que la información sea intuitiva y sin ambigüedades. El sistema de RA hace tracking de la posición y orientación del usuario y superpone, dentro del campo de visión del usuario, gráficos y anotaciones que están alineados con los objetos en el medio ambiente. Este enfoque tiene muchas ventajas ya que la información puede ser presentada de una manera intuitiva y directamente integrada con el medio ambiente [Ses00]. Cuando se implementa correctamente, la RA proporciona al usuario una experiencia de inmersión y permite la interacción entre el mundo virtual y el mundo real. [Mcd03] sostiene que la tecnología de aumentación o ampliación de la percepción humana conocida como RA es una herramienta moderna que están explorando los investigadores. Esta tecnología combina datos virtuales con el ambiente real observado por el usuario por lo que requiere la registración de información virtual con la escena real desde el punto de vista del usuario. [Mil94] introduce un término más genérico, realidad mixta, y lo definen como un *continuum* en donde en uno de los extremos persiste el ambiente real y en el

extremo opuesto persiste el ambiente virtual. La *Realidad Aumentada* y la *Virtualidad Aumentada* constituyen la realidad mixta o intermedia entre estos dos extremos [Ler10].

[Azu97] define tres requerimientos para las aplicaciones de RA:

1. Se deben combinar los aspectos real y virtual
2. La aplicación debe ser en tiempo real e interactiva, y
3. Los objetos deben estar registrados en el mundo tridimensional (3D)

De acuerdo a [Aba11] una aplicación de realidad aumentada tiene las siguientes partes:

- Captura de la escena real: El video capturado puede utilizarse para tracking basado en visión, es decir basado en el análisis de la imagen mediante algoritmos de visión
- Tracking del usuario: puede realizarse mediante dispositivos específicos o puede realizarse tracking basado en visión para lo cual es necesaria la captura de la escena real.
- Generación de la escena virtual: se tiene un mundo virtual, con la información de la posición y orientación del participante se genera una vista acorde del mismo
- Rendering: se combinan las imágenes del mundo real con los objetos virtuales. Los objetos virtuales se *renderizan* y se proyectan en el dispositivo de visualización.

Basándonos en lo detallado anteriormente las líneas de investigación se centraron en 3 grupos: estimación de pose y registración de objetos, clasificación de objetos y renderizado de objetos.

Estimación de Pose y Registración

En términos generales la estimación de la pose (posición y orientación) es un evento en el que se quiere detectar a partir de una serie de puntos dados en un plano, la traslación y rotación de

un objeto en función del movimiento de la cámara entre *frames*. Por otro lado, la registración es tomar estas coordenadas en el plano y proyectar un modelo 3D en éste. El proceso de registro es la superposición de objetos virtuales en una escena real utilizando información extraída de la escena [Pan06]. Como menciona [Aba11] la información virtual tiene que estar vinculada espacialmente al mundo real de manera coherente, lo que se denomina *registro de imágenes (registration, en inglés)*. Por esto se necesita saber en todo momento la posición del usuario, *tracking*, con respecto al mundo real. Cuando se habla de pose (posición y orientación) de un objeto o de una cámara debe existir un sistema de referencia en base al cual se expresan. En RA, al hablar de imágenes registradas significa que tanto las imágenes sintéticas como el mundo real estén en referencia al mismo sistema de coordenadas.

Con el objetivo de ir entendiendo y experimentando con la librería OpenCV, se realizaron pruebas de concepto tales como: procesamiento sobre imágenes, dibujo de líneas, círculos y otros (figura 1). Se realizaron pruebas sobre espacios de colores tales como BGR (Blue, Green, Red), escalas de grises, HSV (Hue, Saturation, Value) y el espacio YUV.

Se desarrolló una aplicación para detectar la presencia de un objeto de color utilizando técnicas de visión por computador. Se realizaron pruebas de *tracking* mediante la implementación de una aplicación de reconocimiento facial (figura 2). Otra prueba en la que se trabajó fue en el calibrado de la cámara. Por otro lado utilizando los datos obtenidos por la calibración, se pudo estimar la pose en un tablero de ajedrez (figura 3). Adicionalmente se evaluó estimar la pose utilizando redes neuronales, más precisamente Redes Convolucionales (*CNN, Convolutional*

Neural Network, en inglés) [Tom14]. Como se verá en la subsección de *Clasificación de objetos*, el uso de Aprendizaje Automático no sólo es útil para clasificar un objeto y a partir de la detección realizar la pose y posterior registración de objetos digitales, sino también se ha experimentado para estimar la pose, por ejemplo, de un cuerpo humano o de otro tipo de objeto.

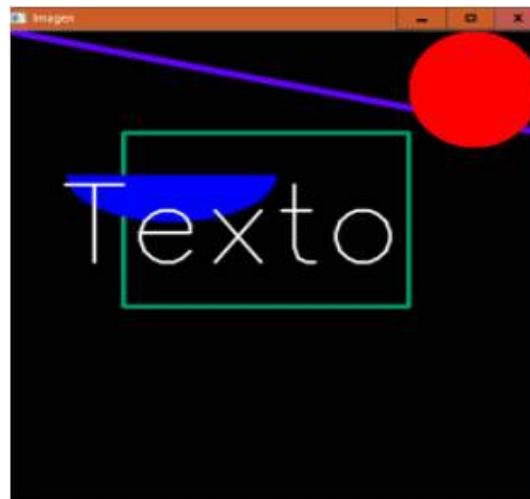


Figura 1

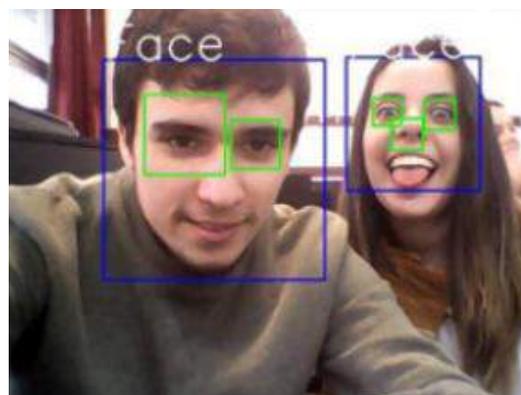


Figura 2

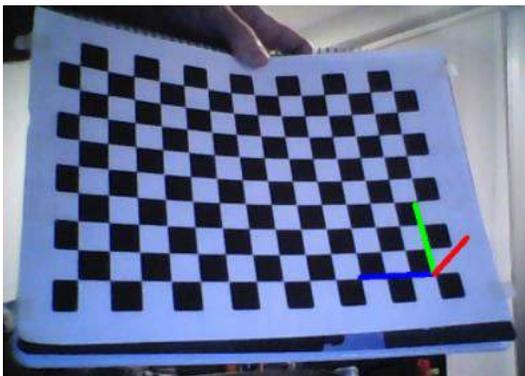


Figura 3

Clasificación de objetos

La acción de clasificar se conoce como “ordenar o dividir un conjunto de elementos en clases a partir de un criterio determinado”¹. Estos criterios pueden definirse según las diferentes características que poseen los objetos a clasificar, por ejemplo los colores, las formas, el tamaño, la textura. Las principales utilidades de los clasificadores son:

- Segmentación de imágenes.
- Reconocimiento de objetos.
- Control de calidad.
- Reconocimiento óptico de caracteres.

Una técnica utilizada para clasificar objetos es “Machine Learning”, o en español “Aprendizaje Automático”. Los sistemas de Aprendizaje Automático se utilizan para identificar objetos en imágenes, reconocimiento de voz, reconocimiento de patrones, etc. Como se menciona en [Yal15], cada vez más, estas aplicaciones hacen uso de una clase de técnicas llamadas “Deep Learning”, o en español “Aprendizaje Profundo”. Para las tareas de clasificación, las capas más altas de representación amplifican aspectos de la entrada que son importantes para la discriminación y suprimen las variaciones irrelevantes. Una imagen, por ejemplo, se presenta en forma de una matriz de píxeles, y las características aprendidas en la primera capa

¹ <https://es.oxforddictionaries.com/definicion/clasificar>

de representación típicamente representan la presencia o ausencia de bordes en una orientación y ubicación particular en la imagen. La segunda capa típicamente detecta arreglos particulares de bordes, independientemente de pequeñas variaciones en las posiciones de los bordes. La tercera capa puede combinar patrones más grandes que corresponden a partes de objetos familiares, y las capas posteriores detectarían objetos como combinaciones de estas partes. Para investigar esta línea se optó por realizar “Estudios exploratorios o Formulativos”. Se dio inicio al proceso de investigación y exploratorio formulando una hipótesis, y a partir de ahí se fue investigando de distintas fuentes, reuniendo la mayor cantidad de información posible para concluir con el objetivo. Las pruebas iniciales fueron ejecutadas sobre *Keras*. *Keras* [Ker18] es una API de Redes Neuronales de alto nivel, escrita en Python y capaz de ejecutarse sobre TensorFlow², CNTK (Microsoft Cognitive Toolkit)³ o Theano⁴. Fue desarrollado con un enfoque que posibilita la experimentación rápida. Al utilizar Keras:

- Permite la creación de prototipos fácil y rápida (a través de la facilidad de uso, la modularidad y la extensibilidad).
- Admite Redes Convolucionales y Redes Recurrentes, así como combinaciones de las dos.
- Se ejecuta sin problemas en la CPU y la GPU.

Renderizado

Se entiende por renderizado (*render*, en inglés) al proceso de generar una imagen, fotorrealista o no, partiendo de un modelo en 2D o 3D. El objetivo principal de esta línea de investigación

² TensorFlow: <https://www.tensorflow.org/>

³ CNTK: <https://docs.microsoft.com/en-us/cognitive-toolkit/>

⁴ Theano: <http://deeplearning.net/software/theano/>

se centró en el análisis de 4 motores de renderizado (incluido motores de video juegos) y cuyo objetivo fue determinar, de acuerdo a un conjunto de parámetros de performance e integración, cuál motor o motores eran los más aptos para integrarlos al framework propuesto. La renderización es el proceso final de creación de la imagen o animación 2D real de la escena preparada. Como menciona [Chu07a], en computación gráfica, el renderizado es el proceso de producir una imagen en la pantalla a partir de la descripción del modelo. Para ello es necesario un motor de renderizado, el cual tiene la función de procesar todas estas imágenes y devolver como resultado el diseño 3D [Chu07b]. Cada uno de estos motores tiene un conjunto de características que hacen que se diferencien unos de otros y delimitan los ámbitos de aplicación de cada uno de ellos. Los 4 motores de renderizado 3D utilizados en la investigación fueron:

- Unity 3D
- Ogre
- OpenScenegraph
- Panda

El objetivo principal de la evaluación consiste en identificar las características que abarca una aplicación 3D; especialmente aquellas que tienen impacto directo en la performance, tiempo de respuesta y fluidez de ejecución de las aplicaciones generadas con el motor [Uni18].

Los parámetros de evaluación fueron agrupados en 2 grupos, por un lado aquellos que están relacionados con la performance, y por el otro, aquellas características que representan otro tipo de ventajas a la hora de integrar dichos motores. Dentro de los parámetros de performance para evaluar en cada motor se seleccionó: velocidad de renderizado y calidad del renderizado. En cuanto a los parámetros de integración se optó por evaluar: lenguajes de programación que soporta, facilidad de

integración, documentación técnica, software libre y ejemplos existentes (cuadro 1). En [Fed17] aportan el criterio de evaluación y experimentación que fueron utilizados para analizar cada uno de los parámetros de evaluación de los motores seleccionados.

	Unity 3D	Ogre	OpenScen	Panda 3D
Velocidad	Medio	Alto	Medio/Bajo	Alto
Calidad	Muy buena	Buena	Buena	Buena
Lenguajes	C#, JavaScript	C#, C++, Java, Python, .NET	C++, Java, .NET, C#, Python, LUA	Python, C++
Integración	Baja	Media	Media	Media
Documentación	Alta	Alta	Media	Alta
Ejemplos	Muy buena	Buena	Buena	Muy buena
Soft. Libre	No	Si	Si	Si

Cuadro 1

Habiendo evaluados los motores de renderizado llegamos a la conclusión que los 2 motores que se adaptarían e integrarían adecuadamente al framework serian Ogre y Panda 3D. Si bien Unity 3D es muy poderoso con suficiente documentación y ejemplos, resultó de difícil integración con nuestro entorno de framework.

Entorno de Desarrollo

Con el fin de estandarizar todas las líneas de investigación se decidió utilizar como sistema operativo Ubuntu 16.04 lts, de modo que se creó y configuró una máquina virtual con el sistema operativo indicado, a través de VirtualBox. La misma se instaló sobre los host Windows 8.1, Windows 10 y MacOS. Utilizar esta distribución resultó útil pues se cuenta con

Python 2.7 instalado junto con las librerías Numpy y Matplotlib, de gran utilidad para poder hacer desarrollos de RA y VA. Por otro lado se instaló de manera separada, OpenCV 3.2.0. Para realizar la investigación y pruebas sobre Deep Learning se descargó el software Anaconda Cloud y se procedió a instalar Keras.

LINEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

- Diseño y desarrollo de un framework de RA y VA multipropósito (aplicado a diferentes disciplinas)
- Aplicaciones de RA y VA utilizando el framework propuesto
- Aplicaciones de Deep Learning aplicables al procesamiento de imágenes y clasificación de objetos
- Modelado y renderizado de objetos 3D utilizando los motores propuestos

RESULTADOS Y OBJETIVOS

A continuación se detallan los resultados obtenidos y los objetivos generales del proyecto:

- Creación del Laboratorio de Procesamiento de imágenes y Visión Artificial. Facultad de Ingeniería y Tecnología Informática. Universidad de Belgrano.
- Curso de Posgrado en Interfaces Avanzadas: Del mundo Real al Virtual⁵. Departamento de Posgrado y Educación Continua. Facultad de Ingeniería y Tecnología Informática. Universidad de Belgrano.
- Diseño y desarrollo del Framework de RA y VA multipropósito. Laboratorio de Procesamiento de imágenes y Visión Artificial. Facultad de Ingeniería y

Tecnología Informática. Universidad de Belgrano.

- Desarrollo de aplicaciones, utilizando el Framework para ser probadas en la Facultad de Ingeniería y Tecnología Informática. Universidad de Belgrano y en empresas privadas de acuerdo a las tratativas que se están gestionando con el sector.
- Aprendizaje de los alumnos en los fundamentos de procesamiento de imágenes, visión artificial y deep learning.
- Aprendizaje del uso de la librería de Procesamiento de Imágenes y Visión por Computador: *OpenCV*.
- Aprendizaje del uso del framework de Deep Learning: *Keras*.

FORMACION DE RECURSOS HUMANOS

La formación de recursos humanos es prioritaria en esta línea de investigación (procesamiento de imágenes y visión artificial) y es por esta razón que se está capacitando a un grupo de alumnos de las carreras de Licenciatura en Sistemas de Información e Ingeniería Informática de la FITI-UB. Uno de los principales objetivos del Laboratorio de Procesamiento de imágenes y Visión Artificial es la formación de recursos altamente calificados en esta temática, pues la intención de la FITI-UB es posicionarse como referente académico en esta área del conocimiento. Adicionalmente con el Curso de posgrado (Interfaces Avanzadas: Del mundo Real al Virtual) se buscará transferir a la comunidad los conocimientos obtenidos en el proceso de construcción del framework, las experiencias obtenidas en dicho proceso y los conocimientos adquiridos en materia de Procesamiento de imágenes y Visión Artificial.

En la actualidad hay en curso una tesis de posgrado, una tesina de grado y capacitación a los alumnos que participan en el proyecto en el marco de esta línea de investigación:

⁵

http://www.ub.edu.ar/sites/default/files/interfaces_avanzadas_final.pdf

- Alejandro Mitaritonna. “Realidad Aumentada para la Identificación de Objetivos Militares” (culminación de tesis de doctorado en la UNLP)
- Silvana Olmedo. “Sistema de Reconocimiento de objetos reciclables” (tesina de grado en Ingeniería Informática en la FITI-UB)
- Carolina Páez, Florencia Vela, Tomás Poeta, Martín Lorenzo y Leonardo Inza (alumnos que forman parte del grupo de investigación del proyecto *Framework multipropósito de Realidad Aumentada y de Visión Artificial*. Laboratorio de Procesamiento de imágenes y Visión Artificial. FITI-UB)

BIBLIOGRAFIA

- [Aba11] M. J. Abásolo Guerrero, C. Manresa Yee, R. Más Sansó y M. Vénere, *Realidad virtual y realidad aumentada. Interfaces avanzadas.*, La Plata, Buenos Aires, Argentina: Editorial de la Universidad Nacional de La Plata (Edulp). 47 N° 380 / La Plata B1900AJP / Buenos Aires, Argentina, 2011.
- [Azu97] R. T. Azuma. *A Survey of Augmented Reality. Teleoperators and Virtual Environments* 6, 4, 355-385, 1997.
- [Chu07a] Chua Hock Chuan, Chang Lanlan, Tan Yap-Peng and Chua Hock-Chuan. *Detection and Removal of Rainbow Effect Artifacts*. IEEE International Conference on Image Processing; San Antonio, Texas, USA. 2007
- [Chu07b] Seong-Ping Chuah, Zhenzhong Chen, Tianxiao Ye, Yap-Peng Tan and Hock-Chuan Chua. *Throughput adaptation for scalable video multicast in wireless networks*. IEEE International Symposium on Circuits and Systems, Taipei, TAIWAN (pp. 1453-1456). 2009
- [Ker18] Keras: The Python Deep Learning library. <https://keras.io/>
- [Ler10] W. Le Roux, *The use of Augmented Reality in Command and Control Situation Awareness*, South Africa: Council for Scientific and Industrial Research, 2010.
- [Mcd03] C. McDonald, *Hand Interaction in Augmented Reality*, Ottawa, Ontario, Canada: The Ottawa-Carleton Institute for Computer Science. School of Computer Science. Carleton University, 2003
- [Mil94] P. Milgram and F. Kishino, “A Taxonomy of Mixed Reality Visual Displays”, *IEICE Transactions on Information Systems*, Vol E77-D, Nr 12, Dec 1994.
- [Mit17] Mitaritonna, A.; Lestani, Juan. *Framework multipropósito de Realidad Aumentada y de Visión Artificial*. Congreso TE&ET, Universidad Nacional de La Matanza, Buenos Aires, Argentina, 2017.
- [Mit15] Mitaritonna, A.; Abásolo Guerrero, M. J. (2015) *Improving Situational Awareness in Military Operations using Augmented Reality*. Proceedings of WSCG 2015. ISBN N°:978-80-86943-72-5, 2015
- [Pan06] Y. Pang, M.L. Yuan, A.Y.C. Nee, S.K. On y Kamal Youcef-Toumi, *A Markerless Registration Method for Augmented Reality based on Affine Properties*, Australian Computer Society, Inc. Seventh Australasian User Interface. 2006.
- [Ses00] S. Sestito, S. Julier, M. Lanzagorta y L. Rosenblum. *Intelligent Filtering for Augmented Reality*. Sydney, Australia, 2000.
- [Tom14] Jonathan Tompson, Arjun Jain, Yann LeCun, Christoph Bregler. *Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation*. Advances in

Neural Information Processing Systems 27.
NIPS, 2014

[Uni18] Optimizing Graphics Performance.
<https://docs.unity3d.com/Manual/OptimizingGraphicsPerformance.html>

[Yal15] Yann LeCun, Yoshua Bengio &
Geoffrey Hinton. Deep learning. Nature volume
521, pages 436–444, 28 May 2015

MEDICIÓN AUTOMÁTICA DE PULSO CARDÍACO UTILIZANDO IMÁGENES DE VIDEO

Eugenia Céspedes¹, Walter Panessi¹, Claudia Ortiz¹

¹ Departamento de Ciencias Básicas, Universidad Nacional de Lujan (UNLu)
eugeniacespedes@gmail.com, wpanessi@unlu.edu.ar, cortiz@unlu.edu.ar

RESUMEN

La división de Ciencias de la Salud y Tecnología del Instituto Tecnológico de Massachusetts, ha propuesto un método para obtener el pulso cardíaco a partir de imágenes de video, aplicando seguimiento de rostro y análisis de componente independiente [1]. Motivados por su descubrimiento, se comenzó a trabajar en una adaptación del método propuesto para generar una metodología de bajo costo y libre de contacto, que posibilite el monitoreo remoto de la frecuencia cardíaca. Para validar los resultados obtenidos, se realizó un análisis comparativo entre las medidas de pulso cardíaco obtenidas en el estudio con las que se obtienen a partir de una medición por fotopletimografía; técnica incluida actualmente en dispositivos de uso cotidiano como los relojes inteligentes y los teléfonos móviles. A partir de allí, se compararon las mediciones de pulso cardíaco previo a la realización de ejercicio físico respecto de las obtenidas post ejercicio físico, con el objetivo de evaluar la adaptación del método y la calidad de los valores obtenidos en estas situaciones concretas, donde los resultados generados deben mostrar diferencias. Los resultados obtenidos nos motivaron a preparar un proyecto de investigación que amplíe los horizontes de este trabajo, buscando mejoras a este proceso y otras aplicaciones del procesamiento de imágenes aplicadas a la salud.

Palabras clave: medición automática de pulso cardíaco, obtención de pulso cardíaco a partir de video, pulso cardíaco sin contacto, análisis

de componente independiente, procesamiento de imágenes.

CONTEXTO

Este trabajo de investigación fue desarrollado por una estudiante avanzada de la carrera Licenciatura en Sistemas de Información con el apoyo de docentes del Departamento de Ciencias Básicas de la Universidad Nacional de Luján. En su comienzo, se desarrolló como un trabajo experimental de fin de curso para la asignatura Procesamiento de Imágenes perteneciente a la carrera. La investigación aborda principalmente la utilización del procesamiento de imágenes aplicado a la salud. Dado los resultados, se está trabajando en un proyecto de investigación a ser presentado en el Departamento de Ciencias Básicas para estudiar la posibilidad de mejorar aún más el proceso y ver otras posibles aplicaciones.

INTRODUCCIÓN

En la actualidad la tecnología se encuentra inmersa en todos los aspectos de la sociedad y es así que la medicina hace uso de los avances tecnológicos para el mejoramiento de las técnicas que se utilizan en el diagnóstico y tratamiento de enfermedades. Según las estadísticas sanitarias mundiales que publica anualmente la Organización Mundial de la Salud, una de las principales causas de mortalidad en el mundo es la cardiopatía isquémica [2], es por esta razón, en conjunto con muchos otros factores, que se requiere de métodos simples y eficientes que permitan un control periódico de parámetros importantes para la detección temprana de esta patología.

En particular existe una variedad de estudios basados en el monitoreo no invasivo de variables fisiológicas. Dentro de estos estudios se pueden encontrar aquellos que realizan adquisición del pulso cardíaco implementando procesamiento de imágenes obtenidas a partir de video. Algunos de ellos lo hacen mediante análisis térmico, estudiando la señal térmica producida por las venas superficiales del cuerpo humano [3, 4], empleando una técnica denominada fotoplestismografía donde se analiza la absorción de la luz en una porción de piel, que varía según el volumen de sangre dentro del tejido [5, 6]. Otros métodos se basan en la obtención de pulso a partir del análisis de sutiles movimientos corporales producidos por la irrigación de sangre en cada pulsación cardíaca [7].

El estudio realizado en este trabajo se enfoca en la obtención de la frecuencia cardíaca a partir del análisis de los canales de color de las imágenes de video [1]. La selección de este método se basa principalmente en el bajo costo computacional de los algoritmos utilizados. En este sentido se ha utilizado un algoritmo para detección de rostro con la finalidad de obtener una zona de interés, se ha utilizado también un algoritmo para la separación ciega de señales por análisis de componente independiente, para la eliminación de ruido y la recuperación de señales no observadas dentro de un conjunto de entrada. En este caso particular la señal a recuperar es la de la frecuencia cardíaca. Lo antes expuesto permite realizar el desarrollo del método, utilizando un ordenador de escritorio, equipado con una cámara web, a diferencia de otros estudios mencionados anteriormente que hacen uso de equipamiento específico como cámaras térmicas o sensores fotoeléctricos.

A lo largo del trabajo se realiza la implementación de una adaptación del modelo de Poh [1]. Se comparan los

resultados con los arrojados por un dispositivo de uso comercial (un reloj inteligente o smartwatch) que utiliza técnicas basadas en análisis fotoeléctrico para obtener las mediciones de pulso cardíaco. Para evaluar la respuesta del modelo y la calidad de los valores obtenidos, los datos se obtienen antes y después de que los participantes hayan realizado ejercicio físico.

LINEAS DE INVESTIGACIÓN Y DESARROLLO

En este trabajo, se realizó una adaptación e implementación del método de Poh [1] para la obtención de la frecuencia cardíaca utilizando elementos de bajo costo y un método no invasivo.

El proceso comienza con la obtención de una filmación de 60 segundos sobre la cual se realiza un proceso de localización del rostro. Esto se hace con un algoritmo basado en el trabajo de Viola y Jones [8]. La obtención del rostro es considerada una zona de interés. A continuación, se realiza un análisis de cada cuadro del video, haciendo una separación ciega de señales por análisis de componente independiente (ICA) [9] para la eliminación de ruido y la recuperación de señales no observadas dentro de un conjunto de entrada.

Con éste método, se analizaron los resultados obtenidos en sujetos de prueba que fueron sometidos a distintas condiciones de actividad física para evaluar la capacidad de la metodología aplicada en condiciones variadas. Se contrastaron estos resultados contra otros métodos de lectura del pulso de manera de validar la correspondencia. En función de estos resultados, se planea realizar a futuro modificaciones en el modelo. En el presente trabajo, por una cuestión de automatización del proceso se seleccionó siempre la segunda componente entre las tres arrojadas como salida, pero ésta no siempre contiene la información relacionada al pulso cardíaco dado que el algoritmo no tiene orden

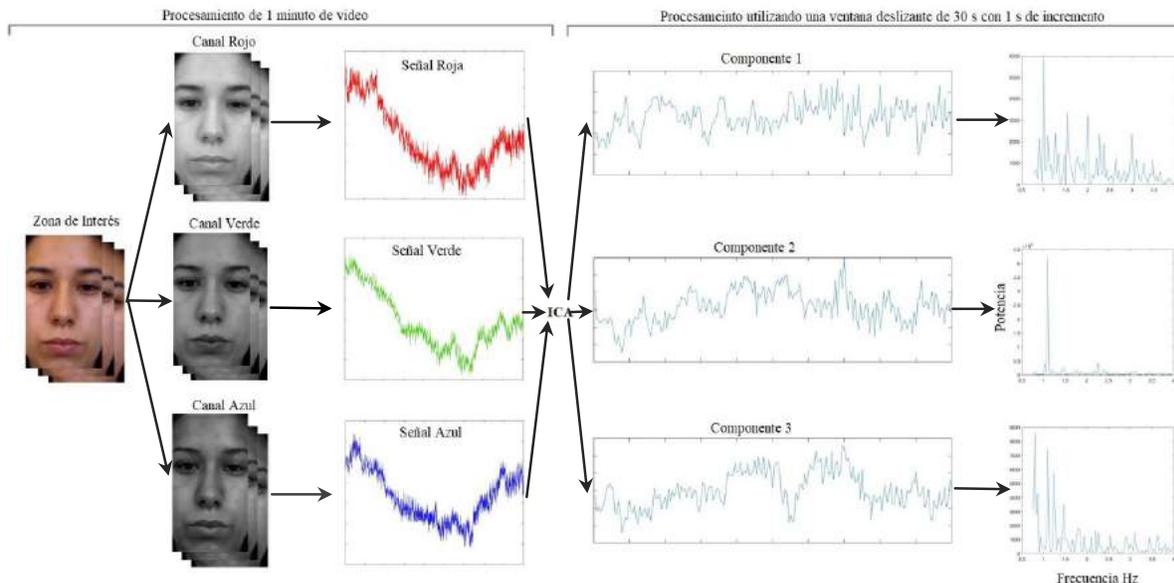


Figura 1. Proceso implementado para la obtención de la frecuencia cardíaca

de salida para los resultados. Se estima que un estudio comparativo entre las salidas para seleccionar la más adecuada, puede arrojar mayores precisiones al modelo.

RESULTADOS OBTENIDOS

En la figura 1 se muestra el proceso utilizado para la obtención del pulso cardíaco, el cual fue desarrollado íntegramente en MATLAB. Los datos de entrada durante la implementación son videos de un minuto de duración en formato mp4 a 30 fps, captados con una resolución de 640 x 480 por una cámara web integrada en una notebook Asus (UX32L), obtenidos de cuatro participantes que se prestaron para el estudio. Sobre cada cuadro de video se realiza detección de rostro utilizando Computer Vision System Toolbox, específicamente “Face Detection and Tracking Using the KLT Algorithm” y se produce una nueva imagen denominada zona de interés (ROI) donde se toma el 60 % del ancho y todo el alto del área generada por las coordenadas que proporciona el algoritmo de detección. Cada ROI se descompone en canales RGB y se realiza un promedio del valor de los píxeles por canal para obtener tres

valores representativos de la imagen, con estos valores obtenidos por cada cuadro que compone el video se construyen tres señales RGB. Luego se aplica sobre cada señal una ventana deslizante de 30 segundos con una superposición del 96.7%, es decir un segundo de incremento, se normalizan las señales RGB utilizando standard score y dado que la absortividad de la hemoglobina difiere en el rango espectral visible e infrarrojo cercano [10], cada sensor de color registra una combinación de la señal que representa el pulso cardíaco, entre otras, es por esto que se obtienen tres componentes independientes subyacentes a las señales de entrada utilizando ICA. Las componentes obtenidas por el algoritmo no tienen un orden de recuperación y es por esto que se les aplica la Transformada Rápida de Fourier (FFT) y se analizan en el dominio de la frecuencia para determinar cuál de ellas representa el pulso cardíaco. El valor del pulso es designado como la frecuencia a la que le corresponde la máxima potencia dentro de la banda de frecuencia operacional, que en este caso es la frecuencia cardíaca (0,74 a 4 Hz o 45 a 240 lpm), utilizando siempre la segunda

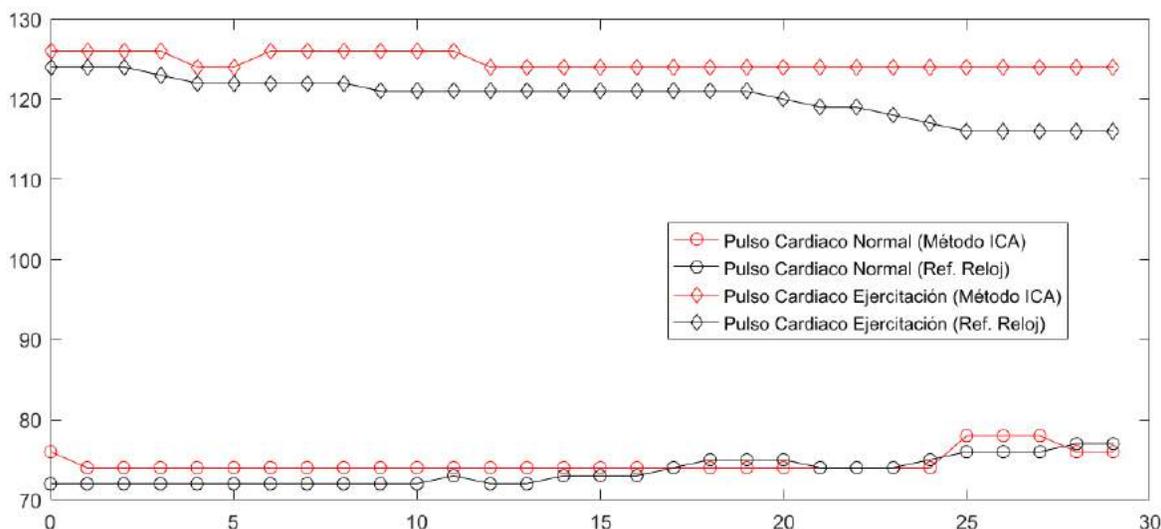


Figura 2: Comparativa de pulso cardíaco obtenido con el método implementado en este estudio y el pulso de referencia obtenido mediante reloj inteligente.

componente para automatización del proceso. Para la validación de resultados se cuenta con un reloj inteligente marca Samsung (Gear 2 Neo) el cual realiza medición de frecuencia cardíaca en tiempo real, se coloca en los sujetos de prueba al momento de la grabación de video y se contrastan los resultados. Para el caso de este estudio se obtiene un primer video donde los participantes se encuentran en reposo y con una frecuencia cardíaca normal y luego se someten a realizar actividad aeróbica intensa en una bicicleta fija por 10 minutos y se procede nuevamente a la grabación de video.

En la figura 2 se pueden observar los resultados obtenidos para uno de los participantes, comparando el pulso arrojado por el método bajo estudio y el obtenido mediante el reloj inteligente, utilizando éste como pulso de referencia. Estos resultados muestran que se obtienen valores satisfactorios en el caso de un pulso normal y también cuando el pulso es elevado por la realización de actividad física. La frecuencia cardíaca cuando el pulso es elevado tiene mayor variabilidad respecto al pulso de

referencia, en comparación con la variabilidad del pulso cardíaco en reposo, esta situación puede atribuirse a que en este caso existe una mayor corrupción de ruido en las imágenes que se atribuye a efectos de movilidad, dado que el sujeto se encuentra con un alto grado de agitación al momento de la grabación producto de la actividad física.

Obtenidos estos resultados puede comprobarse que la adaptación resulta igualmente útil para la obtención de pulso cardíaco cuando el estado del sujeto en estudio varía.

TRABAJOS FUTUROS

Se planifica la continuidad de este trabajo estudiando cómo mejorar el pre-procesamiento de las imágenes intentando eliminar las variaciones de luz y movimiento, las cuales disminuyen la calidad de los resultados obtenidos. Por otro lado, se pretende explorar mejoras de los resultados, definiendo un método de selección para la componente de salida que ofrece el algoritmo de análisis de componente independiente.

FORMACIÓN DE RECURSOS HUMANOS

Se planifica la finalización de la carrera de grado Licenciatura en Sistemas de Información de la alumna Eugenia Cespedes (31 materias aprobadas). La incorporación de docentes auxiliares y estudiantes en actividades de investigación, fomentando la publicación de artículos en Congresos Científicos – Tecnológicos.

BIBLIOGRAFÍA

- [1] Poh, M. Z., McDuff, D. J., Picard, R. W. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics express*, 2010, vol. 18, no 10, p. 10762-10774.
- [2] Página web de la Organización Mundial de la Salud (OMS), “Las 10 principales causas de defunción”. <http://www.who.int/mediacentre/factsheet/fs310/es/> (Consultada: Enero 2017).
- [3] Bourlai, T., Buddharaju, P., Pavlidis, I., Bass, B. On enhancing cardiac pulse measurements through thermal imaging. En *Information Technology and Applications in Biomedicine*, 2009. ITAB 2009. 9th International Conference on. IEEE, 2009. p. 1-4.
- [4] Garbey, M., Sun, N., Merla, A., Pavlidis, I. Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *IEEE transactions on Biomedical Engineering*, 2007, vol. 54, no 8, p. 1418-1426.
- [5] Melchor Rodríguez A., Ramos Castro J. J. Análisis de la Variabilidad de la Frecuencia Cardíaca mediante Fotopleletismografía por imagen. En *Libro de Actas del CASEIB 2014 XXXII Congreso Anual de la Sociedad Española de Ingeniería Biomédica*. 26-28 Noviembre, Barcelona. 2014. p. 1-4.
- [6] Valencia Urbina, C. E. Fotopleletismografía basada en realidad aumentada con aplicaciones al monitoreo funcional en pediatría. 2016. Tesis Doctoral. Universidad Nacional de Cuyo.
- [7] Balakrishnan, G., Durand, F., Guttag, J. Detecting pulse from head motions in video. En *Computer Vision and Pattern Recognition (CVPR)*, 2013 IEEE Conference on. IEEE, 2013. p. 3430-3437.
- [8] Viola, P., Jones, M. Rapid object detection using a boosted cascade of simple features. En *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. IEEE, 2001. p. I-I.
- [9] Hyvärinen, A., Oja, E. Independent component analysis: algorithms and applications. *Neural networks*, 2000, vol. 13, no 4-5, p. 411-430.
- [10] Zijlstra W. G., Buursma A., Meeuwse-Van der Roest W. P. Absorption spectra of human fetal and adult oxyhemoglobin, de-oxyhemoglobin, carboxyhemoglobin, and methemoglobin. *Clinical chemistry*, 1991, vol. 37, no 9, p. 1633-1638.

Aplicaciones de Visión por Computador, Realidad Aumentada y TVDi

María José Abásolo¹², Alejandro Mitaritonna¹³, Sebastián Castañeda¹, Cecilia Sanz¹,
Ramiro Boza¹, Nahuel Prinscich¹, Telmo Silva⁴, Magdalena Rosado¹, Marcelo
Naiouf¹, Patricia Pesado¹, Armando De Giusti¹

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata (UNLP)
{mjabasolo, csanz, mnaiouf, ppesado, degiusti}@lidi.info.unlp.edu.ar

²Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA)

³ Instituto de Investigaciones Científicas y Técnicas para la Defensa (CITEDEF)
amitaritonna@citedef.gob.ar

⁴ Universidad de Aveiro, Aveiro, Portugal

Resumen

La línea de investigación y desarrollo presentada consiste en estudiar, desarrollar y evaluar aplicaciones de Visión por Computador, Realidad Aumentada y Televisión Digital Interactiva. Uno de los principales objetivos es la formación de recursos humanos y fortalecimiento de la investigación mediante el trabajo intergrupar entre diferentes instituciones nacionales y extranjeras.

Palabras Clave: Realidad Aumentada, Visión por Computador, Televisión Digital Interactiva

Contexto

La investigación relacionada con Visión por Computador y Realidad Aumentada forma parte del proyecto 11/F017 "Cómputo Paralelo de Altas Prestaciones. Fundamentos y Evaluación de rendimiento en HPC. Aplicaciones a Sistemas Inteligentes, Simulación y Tratamiento de Imágenes", acreditado por la Universidad Nacional de La Plata (UNLP) en el marco del Programa de Incentivos, continuando con la reciente aprobación del proyecto "Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real" (2018-2021) dirigido por M.Naiouf y "Metodologías, técnicas y herramientas de

Ingeniería de Software en escenarios híbridos. Mejora de proceso" (2018-2021) dirigidos por P.Pesado.

Por su parte, la investigación relacionada con Televisión Digital Interactiva se relaciona con el proyecto REDAUTI Red de Aplicaciones y Usabilidad de la Televisión Digital Interactiva de la convocatoria Redes IX de la Secretaría de Políticas Universitarias del Ministerio de Educación Argentino (2016-2017).

Introducción

Desde hace varios años el grupo de investigación III-LIDI se dedica al desarrollo de aplicaciones en las áreas de Visión por Computador, Realidad Aumentada, Realidad Virtual, Computación Gráfica, Interfaces Avanzadas y Televisión Digital Interactiva (TVDi)[1][2][3][4]. En este artículo se presentan nuestros avances recientes en relación a Visión por Computador en Sistemas de Vigilancia, Realidad Aumentada en el ámbito militar, Realidad Aumentada para brindar información al ciudadano y Televisión Digital Interactiva.

Visión por Computador en Sistemas de Vigilancia

Los sistemas de vigilancia a través de cámaras de video, son útiles para detectar actividades criminales, en aeropuertos, centros comerciales y estacionamientos entre otros. Por este motivo, están en constante desarrollo, siendo uno de los principales objetivos contar

con sistemas de vigilancia automatizados que necesiten poca interacción humana para su funcionamiento y detección de potenciales amenazas. En general, la instalación de cámaras de seguridad no implica un gran gasto económico, pero si la contratación de personas que puedan monitorear esas cámaras durante las 24 horas del día [5].

Es posible adquirir cámaras de seguridad de bajo costo en el mercado local, como por ejemplo la cámara TP-LINK NC220¹. Entre sus varias funciones este tipo de cámaras proveen: detección de movimiento seleccionando un sub área de la zona monitoreada, envío de los *screen shots* por email o a un servidor FTP, detección de sonido y visión nocturna. Por otra parte, el fabricante provee aplicaciones para observar el stream de video desde una computadora o un teléfono móvil. Un problema que presentan estos dispositivos es que la detección de movimiento está basada en las técnicas de sustracción de fondo [6][7] [8] y diferencia entre cuadros [9]. Estos algoritmos son en general muy sensibles a los cambios del entorno tales como la luminosidad o proyección de sombras en el campo visual de la cámara, lo cual produce que muchas veces se detecten falsos positivos. Por otra parte, existen en la bibliografía, diferentes técnicas o algoritmos para detectar personas en imágenes o *stream* de video, entre ellos los histogramas de gradientes orientados [10]. En este algoritmo se extraen descriptores de la imagen y se entrena un clasificador supervisado de manera que dada una imagen es posible determinar si la misma contiene personas. Existen implementaciones disponibles de este algoritmo en la librería OpenCV², como por ejemplo la de Adrián Rosebrock [11] y Dahams [12]. Otro de los problemas que presenta el sistema comercial es que las notificaciones por e-mail no son instantáneas, puede haber un retraso considerable entre la detección de un evento y la recepción del e-mail. Existen por otra parte programas de mensajería instantánea como

Telegram³, el cual es de uso gratuito y provee una API para desarrollar lo que se conocen como *chatbots*.

Realidad Aumentada en el Ámbito Militar

La Realidad Aumentada (RA) se refiere a aplicaciones interactivas en tiempo real donde se visualiza la realidad con elementos sintéticos agregados (objetos 3D, sonidos, texto, etc.) de forma coherente con el punto de vista del usuario.

Habiéndose analizado previamente diferentes proyectos de RA en el ámbito militar, se ha trabajado sobre la implementación del framework denominado RAIOM (Realidad Aumentada para la Identificación de Objetivos Militares) desarrollado en el CITEDEF (Instituto de Investigaciones Científicas y Técnicas para la Defensa) [13] [14][15][16], basado en RA utilizando dispositivos móviles, visión por computador y sensores externos para el reconocimiento, detección, ubicación, identificación y suministro de información contextual. Se diseñaron los componentes de software o módulos que forman parte del *middleware* del framework (figura 1):

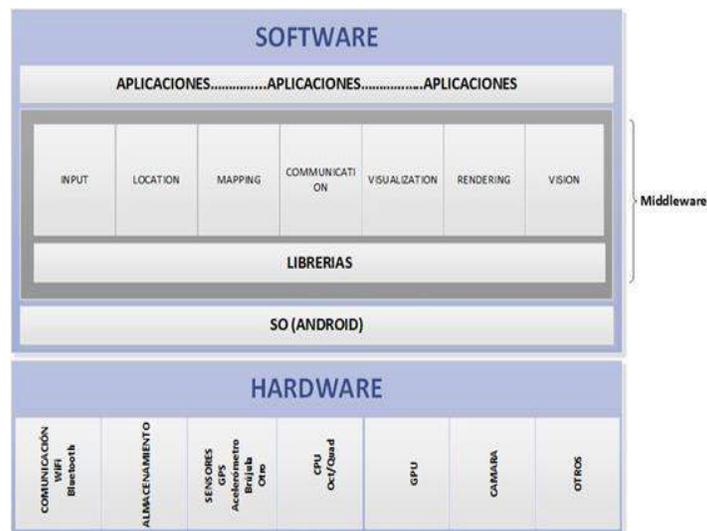


Figura 1. Arquitectura RAIOM

Existen tres módulos del framework que utilizan técnicas de visión por computador, ellos son: el módulo de *Input*, el módulo *Visualización* y el módulo *Vision*. El módulo

¹ http://www.tp-link.es/products/details/cat-19_NC220.html

² <http://opencv.org>

³ <https://telegram.org>

Input permite que se puedan ingresar datos al sistema por medio de voz o gestos. El módulo *Visualization* (figura 2) se encarga del filtrado de información prioritaria corresponde a aquellos datos que el operador puede visualizar como primer capa de información contextual. El módulo *Vision* tiene por objetivo brindar las implementaciones de las funcionalidades necesarias para el reconocimiento de objetos.



Figura 2. Visualización utilizando RAIOM

Se realizó un procesamiento distribuido utilizando los mini board ODROID-XU3. El proceso distribuido planteado está soportado por el diseño de una arquitectura Cliente Servidor. Del lado del Cliente se ejecutan los procesos menos intensivos tales como captura de video, uso de sensores (ubicación, tracking), comunicación, mapeo, y renderizado de imágenes. La implementación de las funcionalidades de reconocimiento de objetos perteneciente al módulo *Vision* fueron desplegadas en los mini board.

Realidad Aumentada para brindar Información al Ciudadano

Algunos autores como M.Lens-Fitzgerald [17] incluyen las aplicaciones basadas en lectura del código de barras y códigos QR como aplicaciones de Realidad Aumentada de Nivel 0, en las cuales los códigos son hiperenlaces a contenidos, sin existir registro de modelos 3D en 3D ni seguimiento de los marcadores.

Dado el uso masivo de *smartphones* este tipo de aplicaciones resulta muy útil a la hora de brindar información contextual al ciudadano. En relación al consumo responsable se hace necesario brindar información clara a la hora de comprar un producto. Es de interés ayudar a fomentar buenas prácticas de consumo que lleven a la comunidad a efectuar un cambio hacia el consumo y desarrollo sostenible y responsable. Actualmente existen distintas aplicaciones que brindan antecedentes sobre móviles para orientar a los consumidores a la hora de realizar sus compra, entre las cuales se citan Think Dirty⁴, EWG'S Healthy Living⁵, GoodGuide Scanner⁶. El problema de dichas aplicaciones es que están orientadas a consumidores residentes en los Estados Unidos, por lo cual, los productos que pueden escanearse son solo los que se comercializan en dicho territorio, y además, toda la información brindada se encuentra en el idioma inglés.

Se detecta la necesidad de desarrollar una aplicación capaz de brindar información, al momento de realizar una compra, sobre los componentes y composición así como el reciclaje post consumo de los productos de consumo diario a partir de la captura con el teléfono celular del código de barra presente en el producto.

Televisión Digital Interactiva (TVDi)

La televisión digital abre la oportunidad de crear aplicaciones interactivas, de ahí la denominación de Televisión Digital Interactiva (TVDi). A través de la TVDi puede ser posible acceder a un conjunto de servicios que abarcan diversos campos como comercio, gestión administrativa, entretenimiento y educación. Esto supone una alternativa al uso de la computadora e internet, lo cual puede facilitar el acceso a la sociedad de la información en sectores que aún no disponen de esa tecnología.

⁴ Think Dirty: <https://www.thinkdirtyapp.com>

⁵ EWG's Healthy Living: <https://www.ewg.org/apps/>

⁶ GoodGuide Scanner:
<https://www.goodguide.com/about/mobile>

Desde el año 2012 se reúnen los investigadores de la RedAUTI para dar lugar a una publicación conjunta para la difusión del estado del arte de la TVDi en Iberoamérica, Portugal y España [18][19][20][21][22][23]. En Argentina en particular, Artola et al [24] presentan una recopilación de aplicaciones de TVDi desarrolladas clasificadas de acuerdo a si están orientadas a ofrecer servicios, a la educación, o a la salud; el tipo de interactividad y el contexto en el que fue desarrollada.

Por una parte, las TIC tienen un amplio potencial de soporte en el campo de la salud, incluyendo sistemas de tele tratamiento que permiten que las personas se mantengan independientes en sus propias casas [25]. Por otra parte, la televisión es una tecnología de gran importancia y gran potencial, uno de los medios preferidos para acceder a la información y al entretenimiento. Dada la proximidad y la familiaridad que los adultos mayores tienen los televisores, el desarrollo de plataformas de TVDi son una forma prometedora de acceder a la información que de otro modo no ser accesible. Como la televisión está presente en la mayoría de los hogares y los adultos mayores interactúan con la misma, existe un desafío de utilizar este dispositivo de interacción en el ámbito de la salud y de asistencia en donde permita ayudar al adulto mayor durante su proceso de envejecimiento.

Líneas de investigación y desarrollo

- Aplicaciones de Visión por computador y Realidad Aumentada para brindar información y servicios al ciudadano
- Aplicaciones de Realidad Aumentada en el ámbito militar
- Contenidos y aplicaciones interactivas para TVDi, en particular dedicados a adultos mayores

Resultados y Objetivos

- Con el objetivo de formar recursos humanos desde el año 2012 se dicta la carrera de postgrado “Especialización en

Computación Gráfica, Imágenes y Visión por Computadora”, Facultad de Informática de la UNLP. (Nº 11.162/12).

- Se ha realizado el dictado de cursos de doctorado relacionados con la temática, como por ejemplo: “Realidad Aumentada”, “Interfaces Avanzadas” y “Tópicos de Procesamiento de Imágenes” (junto a profesores de la Universidad de las Islas Baleares), “Usabilidad” (junto a profesor de la Universidad de Castilla La Mancha). Se organizó para 2018 la visita de profesores de la Universidad Nacional de la Patagonia Austral para el dictado de taller “Dr Nau Herramienta para la Generación

- Coordinación de la RedAUTI, la cual cuenta con la participación de 225 investigadores de 36 grupos de investigación (29 universidades y 7 empresas) de España, Portugal y 11 países latinoamericanos (Argentina, Brasil, Chile, Colombia, Costa Rica, Cuba, Ecuador, Guatemala, Perú, Uruguay, Venezuela), la cual fue creada en 2012 financiada inicialmente por el Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo (CYTED).

- Organización de un evento científico anual denominado Jornadas de difusión y capacitación de Aplicaciones y Usabilidad de la TVDi desde el año 2012, siendo la sexta y última edición jAUTI2017 realizada en Aveiro, Portugal del 12 al 13 de octubre de 2017. Actualmente se está organizando la séptima edición jAUTI2018 a realizarse en noviembre de 2018 en la Universidad Nacional de Quilmes.

- Durante 2017 se realizó la edición conjunta de 3 libros dedicados a Aplicaciones y Usabilidad de la TVDi con aproximadamente 20-25 artículos [26][27][28] y actualmente se trabaja en la edición del último libro.

- Se desarrolló la implementación de un sistema de software de vigilancia que interactúa con un equipo comercial, mejorando por una parte la experiencia de interacción entre el usuario y el equipo y, por otra parte añadiendo niveles de alarmas basados en la detección de intrusos utilizando un algoritmo de reconocimiento de personas

mediante histogramas de gradientes orientados (HOG) en conjunto con una máquina vectorial de soporte lineal (LSVM). Se realizaron test de usabilidad mediante la técnica SUS, para verificar que la funcionalidad agregada al sistema comercial, mejora la usabilidad e interacciones de los usuarios con el sistema. Los resultados obtenidos de estas pruebas confirmaron que la valoración subjetiva de usabilidad de sistemas mejorado fue mayor (puntuación SUS promedio de 95) que la obtenida por el sistema comercial original (puntuación SUS promedio de 57.08).

- Al cabo de 3 años de investigación y pruebas de concepto con el framework RAIOM en desarrollo se han avanzado en la arquitectura cuya etapa es la última de diseño y pruebas. Se han desarrollado un conjunto de aplicaciones de prueba que fueron utilizadas en ambientes controlados y no controlados.

- Se comenzó a desarrollar una aplicación móvil basada en la captura de códigos de barra de productos de supermercado, con el objetivo de brindar información al ciudadano para el consumo consciente y responsable.

- Se está realizando la primera parte de la sobre contenidos y aplicaciones de TVDi para reducir las alteraciones de la marcha en adultos mayores, con el objetivo de probar los desarrollos en Ecuador.

Formación de recursos humanos

La formación de recursos humanos es prioritaria en esta línea, y por esto se implementó la carrera de postgrado mencionada en la sección anterior. En la actualidad hay en curso diferentes tesis de postgrado en el marco de esta línea de investigación:

- Sebastián Castañeda “Sistema de vigilancia para hogares de bajo costo con cámaras fijas mejorado con notificaciones a través de mensajería instantánea y detección de personas en imágenes” Trabajo integrador de la Especialización en Computación Gráfica, Imágenes y Visión por Computadora, Aprobado en 2017

- Alejandro Mitaritonna. “Realidad Aumentada para la Identificación de Objetivos Militares”. Directores: M.J.Abásolo, M. Larrea, F. Simarro (tesis de doctorado en curso)

- Nahuel Prinsich y Ramiro Boza. Aplicación móvil para el consumo consciente y responsable. Directores: Sanz, C. y Abásolo M.J. (tesis de grado en curso)

- Magdalena Rosado “Televisión Digital Interactiva (TVDi) para reducir las alteraciones de la marcha en adultos mayores”. Directores: M.J.Abásolo y T. Silva (tesis de doctorado en curso)

Además, se colabora en la formación de recursos humanos de otras universidades argentinas, entre los cuales se enumeran:

- Ana María Vanesa Sánchez y Mercedes Isabel Castro "Realidad Aumentada en la Fiesta Nacional de los estudiantes" Directores: Pilar Galvez, Nelida Caceres Asesor: M.J.Abásolo (tesis de grado en curso de la Universidad Nacional de Jujuy)

- Lucas Benjamin Cicerchia “Detección de enfermedades y falta de nutrientes en cultivos utilizando algoritmos de Active Learning aplicados al sensado remoto” Directores: Claudia Russo (UNNOBA) , María José Abásolo (tesis doctoral en curso)

- Nahuel A. Mangiarua “Integración escalable de Realidad Aumentada basada en imágenes y rostros” Directores: Jorge S. Ierache (UNLAM), María José Abásolo (UNLP)

Referencias

[1] Abásolo, M.; Mitaritonna, A.; Giacomantone, J.; De Giusti, A.; Naiouf, M.; Manresa, C.; Perales, F.; Sansó Más, R.; Castro, S. *Visión por computador e informática gráfica. Realidad virtual, realidad aumentada e interfaces avanzadas*. WIIC 2013 XV Workshop de Investigadores en Ciencias de la Computación. Universidad Autónoma de Entre Ríos (UADER). 2013. P274 - 279. ISBN 978-987-281-796-1

[2] Abásolo, M.; Mitaritonna, A.; Giacomantone, J.; De Giusti, A.; Naiouf, M.; Perales, F.; Manresa, C.; Vénere, M.; García Bauza, C.. *Realidad Virtual, Realidad Aumentada y TVDI*. WIIC 2014 XVI Workshop De Investigadores en Ciencias de la Computación. Instituto de Desarrollo

Económico e Innovación de la Universidad Nacional de Tierra del Fuego, Antártida e Islas Del Atlántico Sur

- [3] Abásolo, M.; Mitaritonna A.; Encina N.; Vicenzi M.; Borelli L.; De Giusti A.; Naiouf M.; Giacomantone J.. *Realidad Aumentada y Realidad Virtual XVII Workshop de Investigadores en Ciencia de la Computación (WICC 2015)*; Salta, 2015
- [4] Abásolo, M.; Mitaritonna, A.; Bouciguez, M.; Encina, N.; Vicenzi, M.; De Giusti, A.; Naiouf, M.; Giacomantone, J.; Manresa Yee, C. *Realidad Aumentada, Realidad Virtual, Interfaces Avanzadas y Juegos Educativos. XVIII Workshop de Investigadores en Ciencia de la Computación (WICC 2016)*; Concordia, Entre Ríos, 2016 ISBN 978-950-698-377-2
- [5] R. S. Shirbhate, N. D. Mishra, y R. Pande, *Video Surveillance System Using Motion Detection: A Survey*, Advanced Networking and Applications, 2012.
- [6] Z. Zivkovic, *Improved adaptive Gaussian mixture model for background subtraction*, en Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, UK, 26 de agosto, 2004.
- [7] Z. Zivkovic y van der H. F., *Efficient adaptive density estimation per image pixel for the task of background subtraction*, Pattern Recognition Letters, vol. 27, n° 7, pp. 773–780, may 2006.
- [8] P. Kaewtrakulpong y R. Bowden, *An Improved Adaptive Background Mixture Model for Realtime time Tracking with Shadow Detection*, en Proceedings 2nd European Workshop on Advanced Video Based Surveillance Systems, 2001.
- [9] N. Singla, *Motion Detection Based on Frame Difference Method*, International Journal of Information & Computation Technology, 2014.
- [10] N. Dalal y B. Triggs, *Histograms of oriented gradients for human detection*, en IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 2005.
- [11] A. Rosebrock, *Pedestrian Detection OpenCV*, 2015. [En línea]. Disponible en: <http://www.pyimagesearch.com/2015/11/09/pedestrian-detection-opencv>. [Accedido: marzo, 2018].
- [12] C. Dahms, *OpenCV Tutorial 8: Pedestrian Detection using Histogram of Oriented Gradients*. [En línea]. Disponible en: <https://www.youtube.com/watch?v=AKLEuAtFDXQ>. [Accedido: marzo, 2018].
- [13] Mitaritonna, A.; Abásolo, M. J. *Mejorando la conciencia situacional en operaciones militares utilizando la realidad aumentada* (2013) Proceedings of XVIII Congreso Argentino de Ciencias de la Computación. ISBN N°: 978-987-23963-1-2 pp. 356-365.
- [14] Mitaritonna, A.; Abásolo, M. J. (2015) *Improving Situational Awareness in Military Operations using Augmented Reality*. Proceedings of WSCG 2015. ISBN N°:978-80-86943-72-5, 2013.
- [15] Alejandro Mitaritonna, Lucas Pandolfo, Dario Yokhdar y Carlos Esteves. *RAIOM. Introducción a la arquitectura del framework de Realidad Aumentada*. VI Congreso y Workshop Argentino en Ciencias de las Imágenes (ECIMAG 2014)
- [16] Alejandro Mitaritonna, Lucas Pandolfo, Dario Yokhdar y Carlos Esteves. *RAIOM. Introducción a los algoritmos de visión por computador*. VI Congreso y Workshop Argentino en Ciencias de las Imágenes (ECIMAG 2014)
- [17] Lens-Fitzgerald, M. (2009). *Augmented Reality Hype Cycle*. Recuperado de https://www.marketingfacts.nl/berichten/2009042_8_de_augmented_reality_hype_cycle [Accedido: marzo, 2018]
- [18] Abásolo, M.J., De Giusti A. (Eds.) “Anales JAUTI 2012 I Jornadas de Difusión y Capacitación de Aplicaciones y Usabilidad de la Televisión Digital Interactiva” celebrado en La Plata (Argentina) en octubre 2013. ISBN: 978-950-34-0945-9. Universidad Nacional de La Plata (Argentina), 2013. Disponible on-line en <http://sedici.unlp.edu.ar/handle/10915/25926>
- [19] Abásolo, M.J., Castro, C. de (Eds.) “Anales JAUTI 2013 II Jornadas de Difusión y Capacitación de Aplicaciones y Usabilidad de la Televisión Digital Interactiva” celebrado en Córdoba (España) en septiembre 2013, ISBN: 978-84-697-0302-1. Universidad de Córdoba

(España), 2014. Disponible on-line en <http://sedici.unlp.edu.ar/handle/10915/35693>

[20] Abásolo, M.J., Kulesza, R. (Eds.) “Proceedings de jAUTI 2014 III Jornadas de Difusión y Capacitación de Aplicaciones y Usabilidad de la Televisión Digital Interactiva | Workshop WTVI Webmedia 2014” celebrado en Joao Pessoa, Paraíba (Brasil), Noviembre 2014, ISBN: 978-950-34-1188-9, Universidad Nacional de La Plata (Argentina), 2015. Disponible on-line en <http://sedici.unlp.edu.ar/handle/10915/44899>

[21] Abásolo, M.J., Kulesza, R. (Eds.) “Applications and Usability of Interactive TV. Third Iberoamerican Conference, jAUTI 2014, and Third Workshop on Interactive Digital TV, Held as Part of Webmedia 2014, João Pessoa, PB, Brazil, November 18-21, 2014. Revised Selected Papers” ISBN: 978-3-319-22656-9, Communications in Computer and Information Science Springer-Verlag, 2015 <http://www.springer.com/us/book/9783319226562>

[22] A.Bibiloni, F. Perales, M.J.Abásolo (Eds.) “Proceedings del VI Congreso de Televisión Digital Interactiva - jAUTI 2015 IV Jornadas de Difusión y Capacitación de Aplicaciones y Usabilidad de la Televisión Digital Interactiva” celebrado en Palma (España) del 14 al 16 de octubre 2015, ISBN: 978-84-608-5567-5, Universitat de les Illes Balears (España), 2016

[23] Abásolo, María José; Perales López, Francisco; Bibiloni Coll, Antoni (eds.) Applications and Usability of Interactive TV 4th Iberoamerican Conference, jAUTI 2015 and 6th Congress on Interactive Digital TV, CTVDI 2015 Palma de Mallorca, Spain, October 15–16, 2015 Revised Selected Papers. Communications in Computer and Information Science 605, Springer ISBN 978-3-319-38906-6, mayo 2016 <http://www.springer.com/us/book/9783319389066>

[24] Artola, V., Sanz, C., Abásolo, M.J. “Experiencias de TVDI en Argentina” en A.Bibiloni, F. Perales, M.J.Abásolo (Eds.) “Proceedings del VI Congreso de Televisión Digital Interactiva - jAUTI 2015 IV Jornadas de Difusión y Capacitación de Aplicaciones y Usabilidad de la Televisión Digital Interactiva” celebrado en Palma (España) del 14 al 16 de

octubre 2015, ISBN: 978-84-608-5567-5, Universitat de les Illes Balears (España), 2016

[25] Abreu, J. T. F. de, Almeida, P., & Silva, T. E. M. C. da. (2014). Os tele-cuidados e a televisão interactiva. Disponible on-line: https://ria.ua.pt/bitstream/10773/15221/3/2014_In_Envelhecimento%20Saude%20e%20Doen%C3%A7as%20Os%20tele-cuidados%20e%20a%20televis%C3%A3o%20interactiva_done.pdf

[26] Abásolo, María José; Joaquín Pina Amargós (eds.) Libro de aplicaciones y usabilidad de la televisión digital interactiva - V Jornadas Iberoamericanas sobre Aplicaciones y Usabilidad de la Televisión Digital Interactiva, jAUTI2016. Artículos seleccionados Universidad Nacional de La Plata, 2017 ISBN 978-950-34-1514-6 <http://sedici.unlp.edu.ar/handle/10915/61268>

[27] Abásolo, María José; Pedro Almeida; Joaquín Pina Amargós (eds.) Applications and Usability of Interactive TV 5th Iberoamerican Conference, jAUTI 2016 La Habana, Cuba November 21-15, 2016 Revised Selected Papers. Communications in Computer and Information Science 605, Springer ISBN 978-3-319-63321-3, julio 2017 <http://www.springer.com/in/book/9783319633206>

[28] Ferraz de Abreu, Jorge | Abásolo Guerrero, María José | Almeida, Pedro | Silva, Telmo (eds) Anales jAUTI'17 6ª Jornadas de Aplicaciones y Usabilidade de la Televisión Digital Interativa Universidad de Aveiro (Portugal) ISBN 978-972-789-521-2, nov 2017 <http://sedici.unlp.edu.ar/handle/10915/64215>

[ISO10] International Organization for Standardization (2010) *Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems*, 9241-210:2010

[Man12a] Manresa-Yee, C.; Abásolo, M.J.; Mas Sansó, R.; Vénere, M. (2011) *Realidad Virtual, Realidad Aumentada e Interfaces Basadas en Visión*. XV Escuela Internacional de Informática, XVII Congreso Argentino de Ciencia de la Computación CACIC 2011. Editorial EDULP, ISBN 978-950-34-0765-3

Realidad Virtual y Realidad Aumentada como medios para un lenguaje generativo multimodal

N. Jofré, G. Rodríguez, Y. Alvarado, J. Fernández, R. Guerrero

Laboratorio de Computación Gráfica / Dpto. Informática / FCFMyN

Universidad Nacional de San Luis

Ejército de los Andes 950, Tel: 0266 4420823, San Luis, Argentina

{npasineti, gbrodriguez, ymalvarado, jmfer, rag}@unsl.edu.ar

Resumen

El crecimiento exponencial de las nuevas tecnologías prácticamente ha impactado en todos los ámbitos de nuestras vidas. En particular, la Realidad Virtual (RV) y la Realidad Aumentada (RA) son medios que promueven la vivencia de aquellas realidades cuya experimentación activa es imposible. Estas han logrado alterar la percepción de un mundo canónico generando Realidades Alternativas, las cuales han modificado, incluso, la forma en que nos comunicamos.

En este contexto ha surgido una innovación comunicacional multimodal que provee de recursos a los sistemas computacionales para la transmisión de información al usuario en forma rápida, eficiente, natural e intuitiva. En función de ello, se ha focalizado la atención sobre el potencial disruptivo de las modalidades de comunicación provistas por los sistemas de RV y RA favoreciendo las maneras en que las personas pueden ser entrenadas y educadas en relación con la información y las habilidades multimodales específicas que ellas necesitan para resolver problemas.

Esta propuesta de trabajo analiza la configuración de la RV y la RA como sistemas de acceso al conocimiento, con el fin de lograr un lenguaje generativo como complemento motivacional en aquellos procesos que implican un esfuerzo físico y/o cognitivo junto con la creación de entornos multimodales para permitir la adquisición de habilidades y el aprendizaje.

Palabras clave: Realidad Virtual, Realidad Aumentada, Computación Gráfica, Interfaces Humano-Computadoras, Interfaz Natural de Usuario, Comunicación Virtual.

Contexto

La propuesta de trabajo se lleva a cabo dentro del proyecto “Realidades Alternativas como lenguaje generativo aplicado a la solución de problemas reales”. Este proyecto es desarrollado en el ámbito del Laboratorio de Computación Gráfica de la Universidad Nacional de San Luis.

1. Introducción

La ciencia y la tecnología están en un continuo proceso de desarrollo y evolución, lo cual genera grandes transformaciones a nivel social, cultural, educativo, de salud, entre otros; razón por la cual es necesario que los educadores/capacitadores actuales reconozcan, analicen y reflexionen sobre el entorno donde las tecnologías de la información y la comunicación actúan como mediadoras en la formación de las personas [1,2,3].

En los últimos años, la Realidad Virtual y la Realidad Aumentada han despertado el interés de diversas áreas del conocimiento, mostrando su versatilidad y posibilidades como tecnologías innovadoras. Sus capacidades de mostrar escenarios virtuales ficticios o insertar

objetos virtuales en el espacio real, las ha convertido en herramientas muy útiles para presentar determinados contenidos bajo diferentes premisas tales como la educación y la salud [4,5].

El empleo de sistemas de RV y RA constituyen un nuevo enfoque para entrenar, tratar y educar a las personas reforzando el aprendizaje a través de plataformas multimodales [6,7,8,9].

Una plataforma multimodal es un sistema computacional que proporciona una interacción activa (transmisión de información de entrada y salida) con un sistema humano a través de diferentes modalidades sensoriales, correspondientes a los sentidos humanos más desarrollados: la vista, el tacto y el oído [10].

Aunque parezca inconsistente, el *Cómo* nos comunicamos (sobre todo con nosotros mismos) es un factor importante en *Cómo* el entorno se relaciona con nosotros, y es clave en las cosas que hacemos y conseguimos. Desde un punto de vista filosófico se afirma que: "*No sabemos cómo son las cosas. Sólo sabemos cómo las observamos o cómo las interpretamos. Vivimos en mundos interpretativos*" [11]. Es decir, Una persona desarrolla sus ideas, actitudes y manera de pensar a consecuencia del ambiente en el cual se encuentra inmersa. La RV y RA permiten crear Realidades Alternativas diseñadas para conseguir pequeñas metas que implican objetivos, acción, y su consecución [12]. Estas realidades cumplen el rol asociado a un "Coach" en el ámbito deportivo; un director técnico que ayuda a los jugadores para que puedan lograr sus objetivos. El *coaching* es entonces, entrenamiento, el trabajo conjunto del equipo y el coach para elaborar estrategias, corregir defectos y mejorar el rendimiento general [13,14].

Por otro lado, en forma genérica se denomina lenguaje al medio utilizado para la comunicación de información. Desde el punto de vista ontológico, un lenguaje no sólo describe la realidad, sino que por medio de él se genera la realidad. Y más aún, todo lenguaje puede ser utilizado para generar acciones a

través de las cuales se puedan cumplir objetivos, convirtiéndose en lo que se denomina lenguaje generativo [15].

En consecuencia, las Realidades Alternativas creadas por medio de la RV y RA pueden ser consideradas como lenguajes generativos para resolver situaciones que permitan a los usuarios enfrentar una variedad de problemas, modificarlos y/o resolverlos.

Más aún, y desde un punto de vista psicológico, las realidades alternativas pueden ser utilizadas como una herramienta para remover las barreras que hacen que una persona fracase en sus proyectos, convirtiendo el proceso en un aprendizaje transformador.

2. Líneas de Investigación, Desarrollo e Innovación

En función de lo expresado, la Realidad Virtual y la Realidad Aumentada son herramientas poderosas que posibilitarían la creación de nuevos lenguajes de comunicación para la transmisión de información y adquisición de habilidades. Considerando que el ser humano se encuentra dotado de sentidos que le permiten comunicarse/interactuar con el mundo real, es conveniente utilizar como un medio interactivo para comunicarse con una computadora, aquellos estímulos sensoriales que se encuentran fuertemente relacionados con el aprendizaje (visual, táctil y auditivo), [16]. Para ello, los sistemas multimodales a implementar pueden ser clasificados de acuerdo a las capacidades que se pretendan desarrollar o entrenar. Un sistema multimodal enfocado en entrenamiento físico o motor utilizará mecanismos de enseñanza distintos a un sistema que se enfoca en educar en aspectos relacionados a procesos cognitivos [17]. Por lo tanto, de acuerdo al enfoque de enseñanza requerido, se propone distinguir los sistemas multimodales en tres líneas de investigación.

- **Sistemas de repetición:** consisten en comunicar información mediante la reiteración de una tarea o actividad, al mismo tiempo que se incrementa progresivamente la complejidad de dicha tarea. Este tipo de sistemas son útiles en el entrenamiento de procesos mecánicos que usualmente no necesitan ser comprendidos. Entre las aplicaciones usuales se encuentran los sistemas de terapia física, de entrenamiento militar, simuladores de manejo, entre otros [18,19,20].
- **Sistemas de feedback sensorial:** consisten en la comunicación humano-computadora mediante los sentidos (visual, táctil, auditivo). Este tipo de sistemas disminuye el esfuerzo cognitivo requerido para interactuar con la computadora, permitiendo concentrar los esfuerzos en la adquisición de información para la capacitación específica que se está llevando a cabo. En este caso la estimulación sensorial es usada como lenguaje mediado por computadora, el cual actúa como medio para el aprendizaje. Estos sistemas, a su vez, pueden categorizarse principalmente en sistemas *gráficos* (feedback meramente visual), *hápticos* (interacción con el ser humano mediante el sentido del tacto) y *multisensoriales* (más de un sentido como feedback) [21,22].
- **Sistemas motivadores del sujeto:** esta clase de sistemas intentan incentivar el entrenamiento o aprendizaje complementando la capacitación con elementos motivadores que resulten atractivos al usuario. Este es el caso de los *Juegos Basados en Aprendizaje* o *Juegos Educativos*, los cuales en un principio eran netamente visuales, mientras que actualmente incorporan

características tales como *inmersión* y *participación* [23,24].

3. Resultados Obtenidos/Esperados

Las actividades realizadas hasta el momento por este grupo se han enmarcado dentro de un proyecto de investigación de la UNSL, un Proyecto de la Comunidad Europea. y cuatro proyectos de desarrollo tecnológico de la Secretaría de Políticas Universitarias.

Como consecuencia del trabajo elaborado, se ha logrado desarrollar varios sistemas relacionados a RV con interacción multimodal involucrando aspectos verbales y gestuales, así como también se han elaborado aplicaciones de innovación y desarrollo asociadas a la RV y RA.

Actualmente las acciones se encuentran focalizadas en la incorporación de nuevas estrategias que permitan alcanzar una mejor percepción e interacción al mismo tiempo que se logra influenciar al usuario tanto cognitiva como físicamente.

Como resultante, se analizará el impacto logrado por el uso de sistemas de RV y RA en la mejora y adquisición de habilidades funcionales en procesos físicos y cognitivos asociados a las áreas de la salud y la educación, entre otras. De esta manera, se pretende detectar y evaluar la evidencia científica resultante para determinar la envergadura de dichas intervenciones.

4. Formación de Recursos Humanos

Los trabajos realizados han permitido la definición de trabajos finales de carrera de la Licenciatura en Cs. de la Computación (3 finalizados, 1 en ejecución), tesis de Maestría (2 en ejecución y 1 finalizada) y tesis de Doctorado en Ciencias de la Computación (1 en ejecución).

Adicionalmente se ha obtenido una beca de iniciación a la investigación y una beca de perfeccionamiento en investigación otorgadas por la Secretaría de Ciencia y Técnica de la UNSL; y una beca doctoral de CONICET.

5. Bibliografía

1. E. Duval, M. Sharples, y R. Sutherland. *Technology Enhanced Learning: Research Themes*. Springer International Publishing, 2017.
2. R.D. Roscoe, S.D. Craig, y I. Douglas. *End-User Considerations in Educational Technology Design*. *Advances in Educational Technologies and Instructional Design*. IGI Global, 2017.
3. S.Y. Tettegah y M. Gartmeier. *Emotions, Technology, Design, and Learning*. *Emotions and Technology*. Elsevier Science, 2015.
4. T. Jung y M.C. Dieck. *Augmented Reality and Virtual Reality: Empowering Human, Place and Business*. Progressin. Springer International Publishing, 2017.
5. L.T. De Paolis, P. Bourdot, y A. Mongelli. *Augmented Reality, Virtual Reality, and Computer Graphics: 4th International Conference, AVR 2017, Ugento, Italia, Junio 12-15, 2017, Proceedings*. Number pt. 1 in *Lecture Notes in Computer Science*. Springer International Publishing, 2017.
6. Z. Chen, J. Li, Y. Hua, R. Shen, y A. Basu. *Multimodal interaction in augmented reality*. En *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, páginas 206–209, Oct 2017.
7. F. Pallavicini, N. Toniazzi, L. Argenton, L. Aceti, y F. Mantovani. *Developing effective virtual reality training for military forces and emergency operators: from technology to human factors*. En *International Conference on Modeling and Applied Simulation, MAS2015*, páginas 206–210. Dime University of Genoa, 2015.
8. N. Wake, Y. Sano, R. Oya, M. Sumitani, S. Kumagaya, y Y. Kuniyoshi. *Multimodal virtual reality platform for the rehabilitation of phantom limb pain*. En *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, páginas 787–790, Abril 2015.
9. Jorge Bacca et al. *Augmented reality trends in education: a systematic review of research and applications*. *Journal of Educational Technology & Society*, 17(4):133, 2014.
10. I. Wechsung. *An Evaluation Framework for Multimodal Interaction: Determining Quality Aspects and Modality Choice*. *T-Labs Series in Telecommunication Services*. Springer International Publishing, 2014.
11. R. Echeverria. *Ontología del lenguaje*. Distal S R L, 2017.
12. I.R. Management Association. *Virtual and Augmented Reality: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2018.
13. C. Editores et al. *Usos y Perspectivas del Coaching: Un recorrido por el Coaching Iberoamericano actual*. 2015.
14. L. Muradep. *Coaching para la transformación personal: Un modelo integrado de la PNL y la ontología del lenguaje*. Management: Comunicación. Ediciones Granica, S.A., 2012.
15. U. Pinheiro et al. *Ontologia, conhecimento e linguagem: um encontro de filósofos latino-americanos*. FAPERj, 2001.
16. M. Flavin. *Disruptive Technology Enhanced Learning: The Use and Misuse of Digital Technologies in Higher Education*. *Digital Education and Learning*. Palgrave Macmillan UK, 2017.
17. M. Bergamasco, B. Bardy, y D. Gopher. *Skill Training in Multimodal Virtual Environments*, capítulo *Perspectives of Multimodal Virtual Reality (VR) Training Platforms*, páginas 49–68. *Human Factors and Ergonomics*. CRC Press, 2012.

18. Arshad Malik y Anam Zafar. High level activity training through virtual reality in chronic stroke survivor: A case report. *International Journal of Rehabilitation Sciences (IJRS)*, 4(02):36–39, 2017.
19. Nicolas Jofré, Graciela Rodriguez, Yoselie Alvarado, Jaqueline Fernández, y Roberto Guerrero. Virtual trainer for physical activities using a natural user interface. En *XXIII Congreso Argentino de Ciencias de la Computación*, páginas 538–547, 2017.
20. Roger Smith. Military simulations using virtual worlds. Mark Grimshaw (Hg.): *The Oxford Handbook of Virtuality*. Oxford, páginas 666–679, 2014.
21. Nicolas Jofré, Graciela Rodriguez, Yoselie Alvarado, Jaqueline Fernández, y Roberto Guerrero. Gestural interaction for virtual reality environments through data gloves. *Advances in Science, Technology and Engineering Systems Journal*, 2(3), 2017.
22. S. Oviatt et al. *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations*. ACM Books. Association for Computing Machinery and Morgan & Claypool Publishers, 2017.
23. Yoselie Alvarado, Roberto Guerrero, y Francisco Serón. Be civic: An immersive serious game. *EAI Endorsed Transactions on Serious Games*, 18(15), 2017.
24. Minhua Ma, Andreas Oikonomou, and Lakhmi C. Jain. *Serious Games and Edutainment Applications*. Springer Publishing Company, Incorporated, 2011.

Reconocimiento de patrones de imágenes digitales obtenidas mediante microscopio y parametrizadas según la técnica de Micronúcleos y la técnica Ensayo Cometa empleada por el Laboratorio de Citogenética General y Monitoreo Ambiental UNaM-IBS-CONICET para la detección de daños celulares.

Caffetti Yanina Andrea¹, Vera Laceiras María Silvia¹, Acosta Nelson^{2,3}, Pisarello María Inés⁴, Caffetti Jacqueline Diana^{1,5,6}.

¹Facultad de Exactas Químicas y Naturales. Universidad Nacional de Misiones. Posadas, Misiones.

²Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires, Tandil.

³Universidad Nacional de Tres de Febrero, Caseros, Buenos Aires.

⁴Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste. Corrientes.

⁵Facultad de Ciencias Exactas, Físicas y Naturales. Universidad Nacional de Córdoba. Córdoba.

⁶Becaria Postdoctoral CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas), Instituto de Biología Subtropical (UNaM-IBS-CONICET).

yanina007@gmail.com,

vlhsilvia@gmail.com,

el.nelson.acosta@gmail.com,

mainespisarello@exa.unne.edu.ar, jacqui_caffetti@fceqyn.unam.edu.ar

Resumen:

En esta línea de investigación se estudian y analizan diferentes técnicas de clasificación, segmentación y extracción de contornos de imágenes digitales, en particular se centra la atención en las líneas de investigación que utilizan el algoritmo DeepMask, algoritmos evolutivos y redes neuronales. Se presentan aquí algunos resultados preliminares y se pretende aplicarlos a imágenes obtenidas mediante microscopios, específicamente a aquellas parametrizadas según la técnica de Micronúcleos y Ensayo Cometa, ambos empleados por el Laboratorio de Citogenética General y Monitoreo Ambiental de la UNaM-IBS-CONICET para la detección de daños celulares.

Palabras claves: imágenes digitales, microscopio, patrones de imágenes, algoritmos de segmentación, extracción de

contornos, redes neuronales, clasificación de imágenes.

Contexto:

La línea de investigación en que se enmarca el proyecto corresponde a la propuesta de tesis para acceder al título de Magíster en Tecnologías de la Información dictado conjuntamente por la Universidad Nacional de Misiones (UNaM) y la Universidad Nacional del Nordeste (UNNE). Se diseña según lo relevado en el Laboratorio de Citogenética General y Monitoreo Ambiental que se emplaza en la UNaM. El área de Monitoreo Ambiental analiza el daño genético en organismos acuáticos y su impacto sobre la salud humana como consecuencia de la exposición a contaminantes urbanos e industriales presentes en los ríos y arroyos de la Provincia de Misiones. Actualmente se aplican dos técnicas específicas: el Test o Técnica de Micronúcleos y Ensayo Cometa o Electroforesis en gel de células

individuales, a muestras tomadas del Río Paraná, por ejemplo [1]. Para facilitar el trabajo del Laboratorio, se propone investigar la posibilidad de aplicar técnicas de clasificación, segmentación y extracción de contornos de las imágenes digitales producidas por microscopio.

Introducción:

Descripción de actividades del Laboratorio de Citogenética General y Monitoreo Ambiental, Instituto de Biología Subtropical; Facultad de Ciencias Exactas, Químicas y Naturales; Universidad Nacional de Misiones; CONICET (UNAM-IBS-CONICET):

El Laboratorio de Citogenética General y Monitoreo Ambiental inició sus actividades en el año 1989, dedicándose a estudios citogenético-evolutivos en diferentes modelos animales con principal énfasis en peces neo tropicales de agua dulce. Luego se incorporó la línea de citogenética humana, prestando servicios de diagnóstico en convenio entre la UNaM y el Instituto de Previsión Social de la Provincia de Misiones (IPS). Y a partir del año 1993 se incluyó la línea de mutagénesis y monitoreo ambiental, que tiene como objetivos estudiar el impacto de contaminantes en ambientes naturales y en bioensayos de laboratorio a través de técnicas citogenético-moleculares. Actualmente, el laboratorio está constituido por su director, el Dr. Alberto Fenocchio; docentes-investigadores: Mgter. Cristina Pastori, Dra. Jacqueline Caffetti y Lic. Héctor Roncati; así como tesistas, becarios de grado y postgrado: Lic. Ulises Pioli, Lic. Melina Maldonado y los alumnos Angemara Rau y Sergio Müller.

El laboratorio en su línea de investigación

referida a agentes contaminantes acuáticos, aplica técnicas específicas de genética toxicológica, como el Ensayo Cometa o el Test de Micronúcleos. Estos denominados “biomarcadores genéticos” [1], son útiles como señales de alerta temprana en cursos de agua contaminados y, por lo tanto, su análisis resulta de interés predictivo en evaluaciones de estado de calidad de cursos hídricos y manejo de cuencas.

En el test de Micronúcleos, los protocolos aceptados internacionalmente exigen el recuento de 1000 a 2000 células por individuo (considerándose al menos el análisis de 10-15 individuos por cada sitio estudiado). Estas células son clasificadas en distintos subtipos de acuerdo al daño que presentan, por ejemplo: células normales, células con micronúcleos, y otro grupo denominado células con alteraciones de la morfología nuclear que a su vez incluyen: núcleos con lobulaciones, con muescas, con hendiduras, en forma de ocho, con gemaciones, con puentes, binucleadas, entre otras. Este análisis debe hacerse mediante el uso del microscopio, recorriendo todo el preparado en el mayor aumento (100X), según los protocolos de referencia [2, 3, 4].

El Ensayo Cometa, es otra técnica donde se cuentan 100 células por cada individuo (de un total de 10-15 individuos por tratamiento). Estas 100 células se clasifican en 5 clases dependiendo de la intensidad de fluorescencia y el largo de las “colas de los cometas” que son equivalentes a la cantidad de ADN fragmentado o con daño: Clase 0 (sin daño, es decir, no tiene cola); Clase 1 (tamaño de la cola hasta una vez el diámetro de la cabeza); clase 2 (tamaño de la cola hasta dos veces el diámetro de la cabeza); clase 3 (tamaño de la cola hasta

tres veces el diámetro de la cabeza) y clase 4 (casi todo el ADN aparece fragmentado en la cola).

El registro y clasificación de estos tipos celulares se hace manualmente. Una vez se clasifican todas las células en las 5 clases, debe calcularse el score de daño por cada individuo y, a su vez, por cada tratamiento (promediando los datos de los 10-15 individuos).

El registro manual que lleva adelante el Laboratorio, incluye anotaciones en cuadernos del número de células en cada clase y luego pasar los datos a una planilla de Excel para el cálculo del score o ID.

Actualmente cuando se habla de tratamiento de las imágenes digitales, se pueden citar un gran número de técnicas y desarrollos que se encuentran en constante perfeccionamiento, entre ellos la capacidad de segmentación y clasificación de una imagen tratándola como un objeto. La segmentación semántica a nivel de instancia o la segmentación de instancia, por ejemplo, es la tarea de detectar y segmentar conjuntamente instancias individuales de objetos en una imagen [6]. La extracción de contornos en imágenes digitales es una operación que facilita también los procesos de segmentación e identificación de patrones, tanto para tareas de reconocimiento e interpretación, como también de clasificación de objetos [7].

Se inicia la investigación con la posibilidad de obtener imágenes digitales en el Laboratorio de Citogenética General y Monitoreo Ambiental a través del uso del microscopio y una cámara. El objetivo del procesamiento digital de dichas imágenes, que se plantea en esta investigación, es extraer información útil

que pueda parametrizarse según las técnicas de Test de Micronúcleos o Ensayo Cometa.

Líneas de Investigación y Desarrollo:

Para el Test de Micronúcleos se propone investigar la posibilidad de brindar un tratamiento de imagen que logre el corte del recuento al alcanzar el total de células requeridas (1000 o 2000 según corresponda) y a la vez de permitir codificar cada uno de los subtipos celulares en una misma plantilla donde finalmente se calculan tanto los totales de células analizadas en cada clase como las frecuencias finales de cada grupo. Actualmente, las líneas de investigación son: identificación y detección de imágenes a través de la detección de contornos según la propuesta de R. Katz y C. Delrieux [8]. Básicamente, un operador gradiente, herramienta utilizada para la detección de las variaciones en los niveles de intensidad que pudieran corresponder a los contornos de interés. Y, red neuronal convolucional (CNN) única para generar cuadros de delimitación, máscaras de segmentación y clasificar objetos en paralelo, utilizando el algoritmo DeepMask [6]. Investigación en la que una CNN profunda aprende a generar máscaras de segmentación de objetos a partir de una imagen de entrada.

Para la mejora del Ensayo Cometa en test de fluorescencia (captura, clasificación, conteo y gestión de los resultados) se usarán técnicas digitales que combinan algoritmos matemáticos y computacionales que incluyen fuzzy logic [15], redes neuronales [16] y gestión del conocimiento [17]. La captura en este momento se realiza a través de cámaras digitales sobre el microscopio, en un

espacio de total oscuridad para que el reactivo aplicado presente la fluorescencia. Se propone tratamiento de estas imágenes de tipo .jpg, a través de técnicas digitales y/o control predictivo con Modelo MPC para la clasificación [20] y conteo. Con esta clasificación, junto al score y aplicando gestión del conocimiento [17], se intentará ver la evolución de los daños en ADN, o bien, poder predecir compatibilidad en prótesis de acuerdo a aleación de metales en cada individuo. El objetivo del Modelo MPC consiste en el hallazgo de una trayectoria futura de la variable manipulada u. Es decir la implementación de un algoritmo de control predictivo en espacio de estados. Esto se logra utilizando un modelo de predicción que describe el comportamiento de las variables del proceso a controlar, el cual reside en el controlador [18].

Las predicciones dependen de los valores conocidos hasta el instante k_i y de las señales de control futuras. La solución al problema devuelve un vector que contiene las acciones de control futuras cuya dimensión depende del horizonte de control, sin embargo, solamente el primer elemento de este vector debe ser enviado a la planta. Este proceso se vuelve a repetir para cada instante k_i .

Resultados obtenidos/esperados:

Hasta el momento, se estudiaron y analizaron diferentes fuentes de información para definir aquellas técnicas de clasificación de imágenes, incluso aquellas capturadas por microscopio, que tienen menor margen de error en cuanto a la detección de objetos y por consiguiente, menor intervención humana a la hora de procesarlos [9,10,11,12,13,14].

Se espera al finalizar la investigación, el desarrollo de un prototipo basado en la

técnica de manipulación de imágenes que más se adapte a las especificaciones del Laboratorio de Citogenética General y Monitoreo Ambiental de la UNaM-IBS-CONICET en la detección de daños celulares, planteando como línea de investigación futura, la posibilidad de trabajar en otros proyectos que actualmente se llevan a cabo en dicho Laboratorio.

Formación de Recursos Humanos:

Este proyecto de investigación forma parte del desarrollo de dos tesis de posgrado, correspondientes a la carrera de Magíster en Tecnologías de la Información dictada por la Universidad Nacional de Misiones y la Universidad Nacional del Nordeste. Cada tesis tiene como estudio de caso una técnica de análisis utilizada en el Laboratorio, es decir, una tesis trabaja con el Ensayo Cometa y la otra con el Test de Micronúcleos. La tesis que hace referencia al Ensayo Cometa se encuentra bajo la dirección de la Dra. María Inés Pisarello, y la tesis que se vincula al Test de Micronúcleos es dirigida por el Dr. Nelson Acosta. Se articulan con la línea de investigación de la Becaria Postdoctoral del CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas) en el Instituto de Biología Subtropical (UNaM-IBS-CONICET) que lleva adelante la codirección de ambas tesis.

Bibliografía:

[1] GNA. Furnus, JD. Caffetti, EM. García, MF. Benitez, MC. Pastori, AS. Fenocchio. "Baseline micronuclei and nuclear abnormalities frequencies in native fishes from the Paraná River (Argentina)" Brazilian Journal Biology [online]. Vol.74, n.1, pp.217-221. ISSN 1519-6984.

- <http://dx.doi.org/10.1590/1519-6984.13712>, 2014.
- [2] K. Al-Sabti, CD. Metcalfe. “Fish micronuclei for assessing genotoxicity in water”. Mutation Research 343, Netherlands, 1995.
- [3] M. Fenech. “The in vitro micronucleus technique”. Mutation Research 455, Netherlands, 2000.
- [4] J. Barsiene, V. Dedonyte, A. Rybakovas, L. Andreikenaite, OK. Andersen. “Investigation of micronuclei and other nuclear abnormalities in peripheral blood and kidney of marine fish treated with crude oil”. Aquatic Toxicology 78S, 2006.
- [5] JA. Di Rienzo, F. Casanoves, MG. Balzarini, L. Gonzalez, M. Tablada, CW. Robledo. InfoStat versión 2012, FCA, Universidad Nacional de Córdoba, Argentina. [online] <http://www.infostat.com.ar>
- [6] T. Nguyen, A. Shinya, T. Harada, R. Thawonmas. “Segmentation Mask Refinement Using Image Transformations. IEEE Access Dic. 2017.
- [7] A. Jain. Fundamentals of Digital Image Processing. Prentice-Hall, Cambridge, 1996.
- [8] R. Katzy, C. Delrieux. Estrategias Evolutivas para la Detección de Contornos en Imágenes Digitales. Universidad Nacional del Sur, Alem 1253, (8000) Bahía Blanca, 2008.
- [9] M. Datcu and K. Seidd. Image information mining: exploration of image content in large archives. In IEEE Conference on Aerospace, volume 3, pages 253–264, 2000.
- [10] C. Sanz, A. De Giusti. Clasificación de imágenes digitales utilizando patrones N-dimensionales. LIDI. Laboratorio de Investigación y Desarrollo en Informática. Facultad de Informática. UNLP, 2002.
- [11] J. Li, Y. Song, Y. Li, S. Cai, Z. Yang. Automatic Target Segmentation based on Texture for Microscopic Images of Chinese Herbal Powders. Control and Decision Conference (CCDC), pp.1473-1478, at Guiyang, May 2013.
- [12] S. Gupta, S. Sen Purkayastha. Image Enhancement and Analysis of Microscopic Images using various Image Processing Technique. International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 3, pp.044-048, May-Jun 2012.
- [13] R. Gonzales, R. Woods Digital Image Processing. Prentice Hall, New Jersey, 2002.
- [14] Q. Wu, F. Merchant, K. Castleman. Microscopic Image Processing. Pearson, 2010.
- [15] V. Cross; Fuzzy Information Retrieval, [online] <https://link.springer.com/article/10.1007/BF01014019> .1994
- [16] L. Pasa; J. Costa; M. Medeiros; Fusion of Kohonen Maps Ranked by Cluster Validity indexes. Dic 2017.
- [17] H. Takeuchi, I. Nonaka. The Knowledge-Creating Company, How Japanese companies create the dynamics of innovation, Oxford University Press, 1995.
- [18] J. Zambrano, A. González. “Implementación de un algoritmo de control predictivo en espacio de estados sobre una plataforma de simulación desarrollada en Matlab®”. Ingenius. N.º9, (Enero-Junio). pp. 5-14. ISSN: 1390-650X. 2013.

Reconstrucción y animación 3D

Dana Urribarri¹, Juan Ignacio Larregui^{1,2}, Martín Larrea^{1,2}, Silvia Castro^{1,2}

¹Laboratorio de I+D en Visualización y Computación Gráfica (UNS–CIC Prov. Buenos Aires)
Departamento de Ciencias e Ingeniería en Ciencias de la Computación
Universidad Nacional del Sur, Av. Alem 1253, Bahía Blanca

²Instituto de Ciencias e Ingeniería de la Computación (UNS–CONICET)
Departamento de Ciencias e Ingeniería en Ciencias de la Computación
Universidad Nacional del Sur, Av. Alem 1253, Bahía Blanca

{dku, juan.larregui, mll, smc}@cs.uns.edu.ar

RESUMEN

En los campos de la Computación Gráfica y de la Visión por Computadora nos encontramos con dos desafíos importantes. En primer lugar, dentro del área de la Reconstrucción 3D, la recuperación de información de profundidad a partir de imágenes es una tarea laboriosa que requiere no sólo el análisis de características en las imágenes, sino también la correcta aplicación de las propiedades de la geometría de la perspectiva.

Por otro lado, dentro del área de animación, conseguir una animación realista de humanos virtuales no es sólo una tarea sumamente compleja, sino que además, cualquier imperfección es altamente perceptible y produce el rechazo de quien lo observa. Es por esto que muchas aplicaciones utilizan capturas de movimientos para animar sus humanos virtuales.

El objetivo general de esta línea de investigación consiste tanto en el estudio y análisis de técnicas de Reconstrucción 3D a partir de imágenes, como en el análisis de capturas de movimientos para identificar las principales características de los movimientos reales y modelar estos movimientos de manera que permitan ser reproducidos en la animaciones.

Los trabajos se realizan dentro del VyGLab entre becarios y docentes investigadores de la Universidad Nacional del Sur.

Palabras claves: *Computación gráfica, Visión por Computadora, Reconstrucción 3D, Animación 3D, Mo-Cap.*

CONTEXTO

Este trabajo se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) del Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur. Los trabajos realizados bajo esta línea involucran a docentes investigadores y becarios.

La línea de Investigación presentada está inserta en el proyecto acreditado *Análisis Visual de Grandes Conjuntos de Datos* (24/N037), dirigido por la Dra. Silvia Castro y en el proyecto *Análisis de Capturas de Movimientos para la Animación de Humanos Virtuales* (24/ZN33) dirigido por la Dra. Dana Urribarri; ambos financiados por la Secretaría General de Ciencia y Tecnología de la Universidad Nacional del Sur.

1. INTRODUCCIÓN

Dentro de esta línea de Investigación se está trabajando en la reconstrucción 3D a partir de fotografías y en el análisis de Capturas de movimientos (*Mo-Caps*) para la animación de humanos virtuales

1.1 Reconstrucción 3D

La Reconstrucción 3D a partir de imágenes consiste en recuperar la información de la profundidad de la escena en cada uno de los puntos proyectados en el plano de la imagen (píxeles). Se trata de un área que ha recibido gran atención durante el siglo pasado y el

actual, en la cual se ha desarrollado el estudio de las relaciones geométricas entre las proyecciones de una escena a diferentes puntos de vista [17], pero que aún dista de ser un problema solucionado.

En los últimos años, el campo de la Visión por Computadora se ha visto ampliamente afectado por los avances en las técnicas de Deep Learning (DL), especialmente por la aplicación de Redes Neuronales Convolucionales que han obtenido rendimientos superiores a las técnicas tradicionales en la amplia mayoría de los tópicos de interés del campo, como Clasificación de Imágenes, Detección de Objetos, Segmentación Semántica y Reconstrucción 3D, entre otros. La capacidad de estas redes para capturar características de las imágenes de diferentes complejidades sin la necesidad del diseño explícito por parte del humano las hace atractivas para la Reconstrucción 3D, donde la gran variabilidad de los datos visuales y su alta dimensionalidad (potencialmente millones de píxeles) dificulta dicho diseño. Varios trabajos relacionados han sido publicados en los que se entrena las redes para traducir una imagen a un mapa de profundidades [12, 13].

A pesar de su atractivo, una de las principales dificultades en la aplicación del DL a la Reconstrucción 3D se encuentra en la necesidad de contar con un gran volumen de datos etiquetados con información de profundidad para supervisar el entrenamiento de estas redes. La obtención de estas etiquetas debe realizarse mediante sensores y/o en entornos controlados o sintéticos, y en muchos casos requiere un procesamiento manual antes de poder ser utilizadas por las técnicas de DL, haciendo que el volumen de datos no sea fácilmente escalable.

De esto se desprende la necesidad de nuevas técnicas que permitan entrenar redes neuronales convolucionales con menor cantidad de datos, aprovechando el conocimiento geométrico desarrollado durante las últimas décadas, evitando el acercamiento *naive* que deja la totalidad de la tarea de la reconstrucción a las redes neuronales, esperando que no sólo aprendan características de las imágenes, sino

también relaciones propias de la geometría de la perspectiva.

1.2 Animación de Humanos Virtuales

El análisis del movimiento humano (*Human Movement Analysis, HMA*) se refiere al análisis e interpretación de los movimientos humanos en el tiempo [1,3]. Durante décadas, fue un campo de investigación que atravesaba varias áreas: biología, psicología, multimedia, etc. En el campo de la visión por computadora, el *HMA* emergió gracias al video y a la aparición de sofisticados algoritmos de dominio público. Las tecnologías de *Mo-Cap* han agregado al *HMA* la posibilidad de analizar el movimiento a partir de una representación en 3D del esqueleto [14]. Por otro lado, hoy en día, los ambientes sintéticos habitados por humanos virtuales (HHVV) son habituales en un sinnúmero de aplicaciones [6,15,16,18]. Sin embargo, crear un humano virtual (HV) es una tarea sumamente compleja. Dado que estamos acostumbrados a cómo luce hasta el último detalle de un humano, cualquier imperfección en el HV es altamente perceptible y produce el rechazo de quien lo observa [2,7,9]. La teoría del valle inquietante sostiene que cuanto más cerca se está de lograr algo artificialmente humano, mayor es el nivel de rechazo que hay en los observadores humanos [19]. Actualmente existen diversas técnicas para realizar animaciones interactivas en tiempo real [20]; éstas técnicas difieren en el *trade-off* que ofrecen entre la cantidad de control sobre el movimiento, la exactitud y naturalidad del movimiento resultante y el tiempo de cálculo requerido. Elegir la técnica adecuada depende de las necesidades de la aplicación.

Es por esto que en muchas aplicaciones se utilizan *Mo-Caps* almacenados en bases de datos que posteriormente se trasladan a los modelos de HHVV para animarlos. Teniendo en cuenta que se debe almacenar una gran cantidad de *Mo-Caps* para obtener diversidad de movimientos y que estos pueden aplicarse solo en escenarios previamente planeados, surge la necesidad de contar con métodos alternativos para sintetizar humanos que se comporten naturalmente. Una estrategia

tradicionalmente empleada es la utilización de las capturas de movimiento conjuntamente con métodos algorítmicos; sin embargo estos últimos aproximan burdamente las restricciones físicas del cuerpo y del entorno y por lo tanto generan artefactos visuales e intersecciones entre los objetos. El movimiento del cuerpo humano se puede describir desde varios puntos de vista, por ejemplo el mecanismo del movimiento en el espacio y el tiempo, la expresividad cualitativa del movimiento, la trayectoria del movimiento en el espacio, el ritmo y la coordinación del movimiento, entre otras características. Lograr que los HHVV se muevan de manera aceptable es un desafío que requiere identificar las principales características de los movimientos reales y modelar estos movimientos de manera que permitan ser reproducidos en la animación de HHVV.

Un mejor entendimiento de los factores que hacen al movimiento humano reconocible y aceptable es de gran valor en las aplicaciones que requieren realismo en los movimientos de los personajes virtuales [4,7]. La animación realista de un HV es un problema desafiante. Los procesos biomecánicos y fisiológicos que ocasionan el movimiento son difíciles de entender y replicar.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el contexto de esta línea de investigación se están realizando paralelamente los siguientes trabajos:

- Reconstrucción 3D a partir de imágenes en tiempo real.
- Análisis de *Mo-Caps* para la animación de humanos virtuales.

1.1 Reconstrucción 3D

Los desafíos más significativos que se plantean en el proceso de Reconstrucción 3D mediante técnicas de Deep Learning los constituyen:

- La obtención tanto de la secuencia de imágenes como de la información de profundidad de cada píxel.

- La variabilidad de iluminación/color de puntos correspondientes en imágenes desde distintos puntos de vista.
- El entrenamiento de redes neuronales con volúmenes de datos limitados, que tengan la capacidad de predecir correctamente la profundidad de una escena nunca vista (generalización).
- La asignación de información de color a los puntos reconstruidos, pese a cambios de iluminación en las distintas imágenes.
- El diseño y desarrollo eficiente y usable del proceso de reconstrucción 3D en tiempo real.

1.2 Animación de Humanos Virtuales

En el contexto de animación de humanos virtuales, hay varias líneas de trabajo que es necesario atacar para enfrentar este desafío: En primera instancia es necesario contar con herramientas que permitan analizar comparativamente diferentes repeticiones de una secuencia de movimientos. Este análisis puede llevar a identificar secuencias correctamente ejecutadas, medir la experiencia de una persona realizando un movimiento, identificar cuáles son las falencias en la realización de una rutina, etc. Por otro lado, es necesario identificar las propiedades que, además de la trayectoria, hacen al movimiento humano. ¿Qué hace que dos rutinas que ejecutan la misma secuencia de movimientos se perciban de forma diferente? En cuanto a la animación de HHVV, contar con herramientas de comparación permite identificar cuáles son los puntos débiles de los movimientos sintéticos y tomar medidas para corregirlos.

3. RESULTADOS OBTENIDOS/ESPERADOS

1.1 Reconstrucción 3D

Esta línea de investigación explora la reconstrucción 3D de escenas continuas, donde se requiere la estimación de profundidad de cada uno de los cuadros que la constituyen.

El problema puede clasificarse como una regresión densa donde se estima, para cada píxel, su distancia respecto a la cámara. Para ello se han diseñado diferentes arquitecturas de redes neuronales convolucionales, partiendo de redes probadas en tareas de regresión por píxel [10, 11] y se ha supervisado su entrenamiento con diferentes datasets públicamente disponibles [5, 8].

Hasta el momento, la estimación se ha realizado de manera independiente para cada cuadro, sin tener en cuenta la relación de los mismos en el tiempo. Se investigará la aplicación de técnicas de Deep Learning que permitan incorporar esta relación temporal, como la utilización de Redes Neuronales Recurrentes. El objetivo es aprovechar predicciones de cuadros previos en la secuencia de manera de predecir la profundidad de cada cuadro no sólo por la información presente en el mismo, sino también en su contexto.

La correcta estimación de la profundidad en los cuadros de la secuencia representa un paso fundamental en la reconstrucción, modelado y posterior animación de la misma.

1.2 Animación de Humanos Virtuales

Esta línea de investigación se centra en el análisis comparativo de capturas de movimientos en el dominio específico del karate.

En colaboración con el “*Geometry and Graphics Group*” del Departamento de Informática, Bioingeniería, Robótica e Ingeniería en Sistemas (DIBRIS) de la Universidad de Génova (www.unige.it), Italia, se han conseguido *Mo-Caps* de varios karatekas, tanto expertos como novatos, realizando la misma rutina de entrenamiento. Actualmente se logró acondicionar los datos para comenzar con los análisis comparativos:

- Inicialmente, se han completado las capturas de movimiento.
- Luego, las capturas se alinearon y normalizaron en el tiempo. De esta forma el análisis posterior es independiente de la

altura de las personas y de la orientación con la se realiza la rutina.

- Finalmente, se han alineado en el tiempo para evitar que pequeños desajustes en la velocidad de ejecución de la rutina incidan negativamente en la comparación.

A partir de ahora se continuará analizando técnicas para, a partir de parámetros estadísticos, *clusterización* y técnicas de *Deep Learning* como *Recurrent Neural Networks* y *Convolutional Neural Networks* entre otras, comparar las secuencias y distinguir automáticamente las secuencias realizadas por atletas expertos de atletas con niveles de experiencia menores. De esta forma se espera lograr un análisis comparativo de movimientos realizados por atletas expertos, intermedios y novatos.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo que concierne a la formación de recursos humanos se incentiva la incorporación de alumnos que deseen realizar su tesina o trabajo final de carrera en alguno de estos temas. Por otro lado, el Ing. Larregui está realizando su trabajo de tesis doctoral bajo la dirección de la Dra. Castro en el tema de Estimación de Profundidad en tiempo real mediante la utilización de técnicas de Deep Learning.

5. BIBLIOGRAFÍA

- [1] Jasbir Arora and Karim Abdel-Malek. Human Motion Simulation: Predictive Dynamics. Academic Press, 1st edition, 2013.
- [2] James E. Cutting and Lynn T. Kozlowski. Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9:353–356, 1977.
- [3] Yu Ding, Ken Prepin, Jing Huang, Catherine Pelachaud, and Thierry Artières. Laughter animation synthesis. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems, AAMAS '14*, pages 773–780, Richland, SC, 2014. International Foundation for Autonomous Agents and Multiagent Systems.

- [4] Rukun Fan, Songhua Xu, and Weidong Geng. Example-based automatic music-driven conventional dance motion synthesis. *IEEE Transactions on Visualization and Computer Graphics*, 18(3):501–515, 2012.
- [5] Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, vol. 1. 2017..
- [6] S. Hagler, D. Austin, T.L. Hayes, J. Kaye, and M. Pavel. Unobtrusive and ubiquitous inhome monitoring: A methodology for continuous assessment of gait velocity in elders. *IEEE Transactions on Biomedical Engineering*, 57(4):813–820, 2010.
- [7] Ludovic Hoyet, Kenneth Ryall, Katja Zibrek, Hwangpil Park, Jehhee Lee, Jessica Hodgins, and Carol O’Sullivan. Evaluating the distinctiveness and attractiveness of human motions on realistic virtual bodies. *ACM Transactions on Graphics*, 32(6):204:1–204:11, November 2013.
- [8] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The KITTI dataset. *The International Journal of Robotics Research* 32, no. 11 (2013): 1231–1237.
- [9] K. L. Johnson and L. G. Tassinary. Perceiving sex directly and indirectly: Meaning in motion and morphology. *Psychological Science*, 16(11):890–897, 2005.
- [10] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587* (2017).
- [11] Olaf, Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.
- [12] Sudheendra Vijayanarasimhan, Susanna Ricco, Cordelia Schmid, Rahul Sukthankar, and Katerina Fragkiadaki. Sfm-net: Learning of structure and motion from video. *arXiv preprint arXiv:1704.07804* (2017).
- [13] Alex Kendall, Hayk Martirosyan, Saumitro Dasgupta, Peter Henry, Ryan Kennedy, Abraham Bachrach, and Adam Bry. End-to-end learning of geometry and context for deep stereo regression. *CoRR*, vol. abs/1703.04309 (2017).
- [14] Liliana Lo Presti and Marco La Cascia. 3D skeleton-based human action classification. *Pattern Recognition*. 53, C (May 2016), 130–147. 2016.
- [15] Nadia Magnenat-Thalmann and Zerrin Kasap. Virtual humans in serious games. In *Proceedings of the 2009 International Conference on CyberWorlds, CW ’09*, pages 71–79, Washington, DC, USA. IEEE Computer Society. 2009.
- [16] J. Music, M. Cecic, and M. Bonkovic. Testing inertial sensor performance as hands-free human-computer interface. *WSEAS Transactions on Computers*, 8:715–724, April 2009.
- [17] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [18] V. B. Zordan, A. Majkowska, B. Chiu, and M. Fast. Dynamic response for motion capture animation. *ACM Transactions on Graphics*, 24:697–701, 2005.
- [19] Katsu Yamane, Yuka Arika, and Jessica K. Hodgins. Animating non-humanoid characters with human motion data. In Zoran Popovic and Miguel A. Otaduy, editors, *Symposium on Computer Animation*, pages 169–178. Eurographics Association, 2010.
- [20] H. van Welbergen, B.J.H. van Basten, A. Egges, Z.M. Ruttkay, and M.H. Overmars. Real time character animation: A trade-off between naturalness and control. In M. Pauly and G. Greiner, editors, *Eurographics - State-of-the-Art-Report*, pages 45–72, Munich, 2009. Eurographics Association. ISSN: 1017-4656.

Innovación en Educación en Informática

Los Entornos Virtuales Flexibles para el Desarrollo de Competencias en el Área de Algoritmos y Lenguajes de Programación

Myriam G. Llarena, Mario Díaz

Departamento de Informática / Facultad Ciencias Exactas Físicas y Naturales Universidad Nacional de San Juan

Av. Ignacio de la Roza 590 (O). Complejo Universitario "Islas Malvinas". Rivadavia, San Juan, Teléfonos: 4234129, Fax 0264-4234980 Dirección, Teléfonos {myriam.llarena, prof.mariodiaz} @gmail.com

RESUMEN

Este trabajo se encuentra en el marco de un proyecto de investigación que plantea entre sus objetivos aportar conocimiento a la problemática del aprendizaje de la programación en carreras universitarias, en escenarios en los que los límites entre la Educación Formal e Informal se difuminan. Esta investigación propone la generación de estrategias didáctico/pedagógicas, que integrando el entorno institucional de la Educación Formal con el conjunto de aplicaciones y servicios de web 2.0 con las que el alumno se encuentra altamente familiarizado -Educación Informal-, y atendiendo lineamientos de teorías psicosociales, favorezcan el desarrollo de competencias básicas de un alumno que cursa una carrera de programación. Se espera que el conocimiento resultante de la investigación, permita fortalecer la participación del alumno en su proceso de aprendizaje y en la toma de decisiones que implica la construcción de algoritmos computacionales. Se considera que las estrategias propuestas pueden resultar herramientas valiosas para aquellos docentes dispuestos a innovar sus prácticas, mediante la articulación de los espacios de trabajo institucionales con aquellos que el alumno recorre en su cotidianeidad, así como aspectos emocionales-afectivos necesarios.

Palabras clave:

Aprendizaje de Programación. Educación Formal-Informal. Entorno Virtual Flexible. Desarrollo de competencias informáticas.

CONTEXTO

Este trabajo se enmarca dentro del Programa Permanente de Educación a Distancia y del Departamento de Informática de la Facultad de Ciencias Exactas de la Universidad Nacional de San Juan, con aplicaciones en el Campus Virtual de la UNSJ. Integran el grupo de investigadores, docentes de cátedras del área Algoritmos y Lenguajes y Sistemas de Información correspondientes a las carreras Licenciatura en Ciencias de la Computación-LCC- y Licenciatura en Sistemas de Información-LSI- de la Facultad de Ciencias Exactas.

A partir del 2005 el Programa Permanente de Investigación Educación a Distancia desarrolla conocimientos y experiencia en relación a la incidencia de las tecnologías en los procesos de aprendizaje y en la construcción de conocimiento. Posee recursos humanos especializados en Tecnología Educativa con formación en Maestrías y Doctorados en Informática y Educación, así como en Tecnologías Informáticas Aplicadas a Educación.

Además cuenta con profesionales de vasta experiencia en la enseñanza de la programación, que a través de los años han

acumulado un valioso bagaje de conocimiento en la temática.

1. INTRODUCCIÓN

La enseñanza de la Programación no es una tarea sencilla. Criado Clavero (2006) expresa que Programar significa planificar y concretar la secuencia de órdenes concisas que se debe dar al ordenador para la realización de una tarea.

Ciertas habilidades cognitivas son relevantes al momento del aprendizaje de los fundamentos de programación, tales como la capacidad de abstracción, una buena aptitud lógico-matemática y la habilidad para la resolución de problemas de tipo algorítmico.

Esta actividad no es sencilla para los alumnos, lo que produce a través de los años un importante desgranamiento en distintas universidades y en particular en los primeros años de las carreras LCC y LSI de la Facultad de Ciencias Exactas de la UNSJ. Esto es atribuido a que en general no han adquirido, antes de ingresar a la universidad, algunas competencias necesarias tal como capacidad de abstracción, análisis, síntesis; responsabilidad y compromiso en su proceso de aprendizaje; actitudes para trabajar en grupo y para comunicar sus puntos de vista en forma oral y escrita, todas ellas fundamentales para la construcción de algoritmos.

Tejada Díaz(2017) puntualiza que promover en las universidades un proceso de formación para posibilitar el acceso al conocimiento y disminuir las brechas cognitivas, tecnológicas y sociales, implica mejorar las estrategias para facilitar un aprendizaje integral. Es necesario modificar y transformar prácticas relacionadas con el enseñar y aprender, de manera que:

- De un proceso centrado en la enseñanza se pase a uno centrado en el aprendizaje.

- Se combinen contextos y espacios físicos presenciales con los ambientes virtuales, se utilicen redes que favorezcan la generación del conocimiento colaborativo, la socialización, el intercambio, donde los estudiantes se convierten en sujetos que aprenden y enseñan.

- Se logre la integración sistémica y holística expresada en competencias, como cualidad humana que se configura como la máxima expresión de logros de aprendizajes y de desarrollo socio profesional.

Por su parte, Dolors Reig (2017) expresa que es necesario reflexionar sobre la educación en un contexto en el que la tecnología se hace ubicua y permea todos los ámbitos de la vida de los jóvenes. La pregunta ya no es cómo se puede aprender mejor con las TIC, sino como están influyendo en el aprendizaje, para valorar sus riesgos, pero, sobre todo, las grandes oportunidades que ofrecen. El auto-aprendizaje, los nuevos contextos educativos y los aprendizajes no formales que las TIC facilitan y potencian deben ser aspectos a tener muy en cuenta por parte de profesores. Los cambios tecnológicos producidos por la digitalización de la cultura producen nuevas exigencias a las instituciones educativas en términos de la construcción de conocimiento y propone pensar nuevos escenarios de aprendizaje

Esa reflexión debe hacerse desde los distintos componentes de la educación:

- El desarrollo cognitivo (secuenciación, esquemas espacio-temporales, estructuras cognitivas...).
- La individualización (lo emocional-afectivo, la consistencia ética y moral...).
- La socialización (como persona y como ciudadano).

Joyanes Aguilar (1996) indica que "para la mayoría de los estudiantes, el conocimiento inicial eficiente de conceptos básicos tales como algoritmos y estructura de datos, así como el modo de aprender a resolver problemas mediante computadoras, suele ser vital debido a la trascendencia que un aprendizaje gradual y correcto supondrá para su carrera y sobre todo para sus primeros cursos en la universidad".

Es importante no perder de vista esta afirmación que permanece vigente en la actualidad. La construcción de este modelo de aprendizaje debe estar nutrida de las ventajas que aportan la evolución del conocimiento y las nuevas herramientas que aportan las teorías de vanguardia.

Disminuir brechas cognitivas implica atender las características de alumnos que se desarrollan en un contexto impactado fuertemente por diferentes servicios y aplicaciones de la web 2.0, que les permite dejar de ser un consumidor pasivo y transformarse en productor de contenidos, que comparte en diferentes redes sociales.

Según Besada Estevez (2014), en este contexto, los límites tradicionales que separan la Educación Formal de la Educación Informal se han vuelto permeables, en función del repertorio de contenidos digitalizados disponibles a todos los usuarios: el acceso al conocimiento es cada vez más sencillo, los procesos de enseñanza/aprendizaje se dan cada vez en más contextos y de una manera más dinámica en una fórmula próxima al aprendizaje bajo demanda.

Para esta investigación se retoma la conceptualización de Coombs en Sirvent (2006) que define la Educación Formal como "altamente institucionalizada, cronológicamente graduada y jerárquicamente estructurada". La Educación Informal es aquella que se realiza durante toda la vida, "en la que se adquieren

y acumulan conocimientos, habilidades, actitudes y modos de discernimiento mediante las experiencias diarias.

Por tanto el gran desafío de las instituciones de educación superior es desarrollar ecosistemas de aprendizaje (García-Holgado & García-Penalvo, 2013; Llorens-Largo et al., 2014), donde se integren tanto las herramientas y recursos institucionales, como aquellos usados por los estudiantes en su aprendizaje personal. Esto requiere entonces reflexionar sobre los espacios personales donde el estudiante desarrolla y alcanza el conocimiento; espacios que se conocen en el campo científico y académico como Entornos Personales de Aprendizaje (PLE) (Attwell, 2007; Adell & Castaneda, 2010; Casquero, Portillo & Benito, 2013).

Estos nuevos ecosistemas, a los que llamamos Entornos Virtuales Flexibles (EVF) favorecen la consideración de la dimensión individual, las características únicas de cada alumno, con sus propias limitaciones, fortalezas, intereses, capacidades y estilos de aprendizaje. La integración de ambos espacios aparece como una alternativa válida para superar los limitantes encontrados en el uso de las plataformas institucionales en la educación superior (Humanante-Ramos, García-Penalvo, & Conde, 2015).

Por otro lado, frente a enfoques tradicionales de la formación basada en el conocimiento, en la actualidad se propone el enfoque de la formación basada en la competencia (FBC). En el aprendizaje por competencias, el estudiante necesita ser capaz de manejar el conocimiento, actualizarlo, seleccionar la información, conocer las fuentes de información y comprender lo aprendido para integrarlo a su base de conocimiento y adaptarlo a nuevas situaciones. (Blas Aritio, 2007).

Atender la componente emocional y afectiva desde lo individual y social, expresada por

Dolors Reig, requiere investigar y analizar los aportes que la ciencia realiza sobre los sistemas atencionales, los aspectos a tener en cuenta para mejorar el ambiente de aprendizaje, permitiendo comprender y mejorar el estado emocional, la autoestima, la motivación, la adaptabilidad, la empatía, la toma de decisiones, la cooperación, el trabajo en equipo (Logatt Grabner, 2015). Estas teorías consideran la inteligencia como una capacidad (Howard Gardner) en donde se relacionan la genética y el medio ambiente (Logatt Grabner, 2015) y por tanto puede desarrollarse. Gagné (1986) explica que las actividades internas que intervienen en el proceso de aprendizaje tienen una estrecha conexión con las actividades externas, que da lugar a determinados resultados de aprendizaje. Es necesario por tanto que el docente genere condiciones externas lo más favorables posibles para que las fases señaladas por el autor: motivación, comprensión, adquisición, retención, generalización, ejecución y realimentación, sean transitadas sin dificultades a la hora de construir algoritmos computacionales.

En este escenario planteamos los interrogantes: ¿Es posible definir estrategias que permitan acercar las componentes formal e informal que ofrece la tecnología para favorecer el aprendizaje de programación? ¿Podrán proponerse nuevas estrategias pedagógicas que permitan el crecimiento individual y grupal de nuestros alumnos a partir del conocimiento propiciado desde teorías que aborden aspectos socio-emocionales? ¿Favorecerán estas estrategias el desarrollo de las competencias específicas necesarias para un estudiante de programación que ayuden a disminuir el desgranamiento? ¿Qué indicadores son adecuados para validar la implementación de estas estrategias?

Esta investigación se propone favorecer el desarrollo de competencias básicas de un alumno que cursa una carrera de

programación a partir de estrategias didáctico pedagógicas que integren al entorno institucional de la Educación Formal, aplicaciones y servicios de web 2.0 con las que el alumno se encuentra altamente familiarizado y atendiendo lineamientos de teorías psicosociales.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Los ejes a abordar en esta investigación son, el conocimiento tecnológico, pedagógico, disciplinar y de teorías psicosociales, dado que integrar las TIC en el proceso de enseñanza/aprendizaje implica no solamente conocer las herramientas, sino también “reacomodar” las prácticas educativas, revisar y resignificar los conocimientos pedagógicos y disciplinares.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Se han planteado los siguientes objetivos a alcanzar:

- Identificar las competencias específicas para el aprendizaje de la Programación.
- Identificar obstáculos que presentan los alumnos que inician la tarea de programar.
- Definir estrategias de uso de un EVF para el desarrollo de competencias en el aprendizaje de la Programación.
- Aplicar principios básicos propiciados desde teorías psicosociales en las estrategias pedagógicas.
- Evaluar las estrategias propuestas en cátedras de carreras LCC y LSI-FCEfyN.
- Conformar una comunidad virtual entre educadores del área Algoritmos y Lenguajes de Programación, en relación a la problemática de entornos virtuales flexibles y enseñanza aprendizaje de programación.

4. FORMACIÓN DE RECURSOS HUMANOS

Se considera que la capacitación de los docentes de las distintas áreas de las carreras LCC y LSI para la integración de las herramientas usadas por los alumnos en su aprendizaje de la vida cotidiana, a la educación formal y la atención de la componente emocional permitirá disminuir el desgranamiento producido a través de los años.

Miembros de este equipo son docentes de la Diplomatura “Educación y Nuevas Tecnologías en tiempos de convergencia”- FCEFyN- por lo que estas estrategias podrían adecuarse a otras áreas disciplinares.

Actualmente se realiza asesoramiento a tesis de Maestría y de Doctorado en temáticas afines a las abordadas en este proyecto.

5. BIBLIOGRAFIA

Adell, J., Castañeda (2015) *Las pedagogías escolares emergentes*. Cuadernos de Pedagogía N° 462, Sección Monográfico, España: Editorial WoltersKluwer.

Alfin-EEES (2014) *Habilidades y competencias de gestión de la información para aprender a aprender en el Marco del Espacio Europeo de Educación Superior*. <http://www.mariapinto.es/alfineees/AlfinEEES.htm>

Atwel, G (2007) *Personal Learning Environments: the future of learning? eLearning Papers*, volumen(2(1)), pp(1-7). <http://www.elearningeuropa.info/files/media/media11561.pdf>

Blas Aritio, Francisco (2007) *La formación profesional basada en la Competencia*. Revista de la Asociación de Educación de España, volumen (7).

Besada Estevez Alba (2014) *Las TIC y su impacto en la sociedad*. <http://stellae.usc.es/red/blog/view/111106/la-s-tic-y-su-impacto-en-la-sociedad>

Criado Clavero, M.A. (2006) *Programación en Lenguajes Estructurados*. Ed. Alfaomega, España, pp(8).

Humanante-Ramos, Garcia-Penalvo, F. J., Conde, M. A. (2015b) *Personal learning environments and online classrooms: An experience with University students*. Revista Iberoamericana de Tecnologías Del Aprendizaje, volumen(10(1)), pp. (26-32)

Joyanes Aguilar, Luis (1996) *Fundamentos de Programación. Algoritmos y estructura de datos*. Ed. McGraw-Hill/Interamericana, España, pp (4).

Reig Dolors (2017) *Blog El Caparazón referido al ámbito de la innovación, la educación y la tecnología* <http://www.dreig.eu/caparazon/>

Tejada Díaz Rafael (2017) *Paradigmas Prospectivos de los contextos Universitarios*. <http://redesib.formacionib.org/blog/paradigmas-prospectivos-de-los-contextos-universitarios>

Sirvent, M. T. y Otros (2006) “ *Revisión del Concepto de educación No Formal, Cuadernillo de la Cátedra: Educación No Formal: Modelos y Teorías*. Bs.As. Facultad de Filosofía y Letras de la UBA.

Sistema Experto con actualización de reglas en cascada para la construcción de indicadores en Prácticas de Educación Digital y Robótica Educativa.

Ing. Solange Schelske
Universidad Nacional de Misiones
Posadas, Argentina
solange.schelske@gmail.com

Dr. Nelson Acosta
Universidad Nacional del Centro de la Provincia de Buenos Aires
Universidad Nacional de Tres de Febrero
Buenos Aires, Argentina
el.nelson.acosta@gmail.com

RESUMEN.

El presente trabajo se centra en la propuesta de construcción de una matriz de indicadores de competencias y habilidades en tiempo real a partir de la implementación de dos sistemas expertos vinculados en cascada, con una matriz inicial de indicadores de base teórica para luego del aprendizaje del segundo sistema experto construir una matriz de indicadores de base empírica de prácticas de educación digital y robótica educativa.

Palabras clave: Robótica educativa, Sistemas Expertos, Inteligencia Artificial.

CONTEXTO.

La tecnología es una materia multidisciplinar, en la que los alumnos no sólo adquieren conocimientos en tecnología y ciencia, sino que también desarrollan habilidades creativas y sociales [1].

La tecnología educativa es un aprendizaje basado en la resolución de problemas aplicado a proyectos reales en los que alumnos y alumnas realizan los pasos necesarios para crear una herramienta o a resolver un problema. La robótica educativa es un sistema de enseñanza interdisciplinaria que potencia el desarrollo de habilidades y

competencias en los alumnos.

En este sentido, es fundamental integrar el aprendizaje de contenidos con la creatividad, innovación y el aprendizaje basado en proyectos ya que cumplen un rol significativo en las tendencias de incorporación de conocimientos tanto en la programación como en la robótica [2].

El surgimiento de la inteligencia artificial como una disciplina científica, y el desarrollo tecnológico que se ha impulsado en el campo de los sistemas expertos, ha abierto una nueva gama de posibilidades a docentes y alumnos en el marco de un modelo educativo centrado en el aprendizaje y no en la enseñanza [3].

La importancia de la construcción de indicadores reales de competencias y habilidades digitales según las edades de los estudiantes durante el proceso de aprendizaje, los conceptos fundamentales de los sistemas expertos y la actualización de reglas en cascada, los beneficios y limitaciones que ofrecen en el campo educativo y aplicaciones en el área de la robótica educativa constituyen temas de estudio en el presente trabajo.

1. INTRODUCCIÓN.

La Inteligencia Artificial (IA) tiene como propósito reproducir las acciones y el razonamiento de los seres vivos inteligentes en dispositivos artificiales, con el objetivo de conseguir una teoría comprensiva de la inteligencia, tal como aparece en animales y máquinas. El ejemplo más paradigmático de máquinas inteligentes, son los Sistemas Basados en Conocimientos (SSBBCC), y los Sistemas Expertos (SSEE), como puede verse en [Britos, 2001; Rossi, 2001; Rizzi, 2001; Bermejo, 2002; Ierache, 2001; Cao, 2003; Diez, 2003; Gómez, 2003; Hossian, 2003]. A la actividad de construir estos sistemas se la denomina Ingeniería del Conocimiento [Hayes-Roth et al, 1983]. Un sistema experto se refiere a un software que permite imitar el comportamiento de un experto humano en la solución de un problema en particular [Hu, 87].

Durante años el aprendizaje se basa en un modelo de enseñanza secuencial, donde primero se trabaja la teoría para luego aplicar en la práctica (y no en todas las asignaturas). Cuando incluimos herramientas digitales o TIC en los procesos de enseñanza, el modelo se invierte, aprender la práctica para luego entender la teoría.

Las asignaturas de ciencias son más propensas a no quedarse en la teoría, si bien es verdad que en la actualidad el mundo atraviesa una crisis respecto a cuáles son los modelos de aprendizaje con mayor impacto para el contexto en el que nos encontramos, se necesitan más herramientas para llevar adelante este tipo de actividades.

Cuando se trabaja con programación o robótica se fortalecen los enlaces en cuanto al trabajo en equipo y la colaboración entre pares; es por ello que se proponen metodologías ágiles de proyectos. De este modo se busca provocar en los estudiantes, entusiasmo por desarrollar habilidades y competencias que les permitan la construcción de saberes.

Los avances de la inteligencia artificial en el campo de los sistemas expertos, están otorgando a profesores y estudiantes nuevas oportunidades, donde el proceso educativo,

respetando la diversidad, permite la adquisición de competencias que han sido difíciles de alcanzar en la educación tradicional, tales como: aprendizaje auto dirigido, gestión del propio conocimiento, automotivación y autodirección [3].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO.

El desarrollo de la propuesta comienza con el planteamiento de la situación real en base a una matriz de indicadores de competencias y habilidades digitales conformando un modelo inicial según las edades de los estudiantes. A partir de la matriz generada, se presenta la construcción del modelo para la representación del conocimiento experto en robótica educativa (integrando disciplinas como programación, psicopedagogía, diseño industrial, electrónica), y se diseña y formula el sistema experto. Este simula los procesos de reflexión propios de los expertos humanos en las áreas anteriormente mencionadas.

Se propone el desarrollo de una plataforma web para la interacción con el estudiante con la implementación de la técnica de gamificación, donde por medio de desafíos se realice el registro de las actividades y los avances en la incorporación de los conocimientos.

Cada desafío estará vinculado a un conjunto de competencias y habilidades, lo que nos permitirá obtener los avances.

El registro de cada estudiante será la base de hechos del sistema experto, quien realizará las comparaciones necesarias para obtener el informe de acuerdo a los indicadores alojados en la base de conocimiento.

En el caso de que el resultado sea positivo, y los avances en los desafíos correspondan a los esperados según los indicadores, el desafío será catalogado como buena práctica y conformará una base de casos de éxitos.

En los casos donde los resultados no correspondan a ningún caso dentro de la base de reglas del sistema experto, será enviado a otro sistema experto de casos excepcionales, para observar el comportamiento de este tipo

de casos e identificar si son cambios en los indicadores o bien casos excepcionales.

3. RESULTADOS OBTENIDOS/ESPERADOS.

Uno de los resultados esperados es que el sistema experto oriente al profesor en las decisiones metodológicas más eficaces para su práctica educativa teniendo en cuenta los diagnósticos elaborados sobre las características de los estudiantes del entorno de enseñanza-aprendizaje y las habilidades y destrezas del profesor.

Una vez implementado y en funcionamiento el sistema experto tendrá la capacidad de identificar casos excepcionales a las reglas ingresadas por los expertos humanos. El sistema experto identificará en una base de datos temporal estos casos excepcionales para determinar cuándo se convierta en un nuevo patrón o en una modificación real de los indicadores.

Es por ello que se plantea un segundo sistema experto que se actualice a partir de estos casos excepcionales, creando nuevos indicadores, adaptando así a los procesos de enseñanza reales de los estudiantes.

4. FORMACIÓN DE RECURSOS HUMANOS.

El presente trabajo representa la línea de Trabajo Final de Maestría correspondiente a la Maestría en Tecnologías de Información.

La implementación plantea un impacto en estudiantes de los trayectos de Aprendiendo con Robótica de 5 a 20 años de la Escuela de Robótica de Misiones, realizado en 4 etapas según cada trayecto y matriz de indicadores.

5. BIBLIOGRAFÍA.

- [1] «Camp Tecnológico,» 14 Junio 2017. [En línea]. Available: <https://camptecnologico.com/implantacion-curricular/>.
- [2] J. M. García, «Robótica Educativa. La programación como parte de un proceso,» RED-Revista de Educación a Distancia, 2015.
- [3] E. V. Quesada, «sistemas expertos para la enseñanza y el aprendizaje de la

matemática en la educación superior,» cuadernos de investigación y formación en educación matemática, vol. 2, nº 3, pp. 45-67, 2007.

[4] S. Yazyi, «Una experiencia práctica de Scrum a través del aprendizaje basado en proyectos mediado por TIC en un equipo distribuido,» Trabajo de Fin de Master en las TIC en Educación de la Universidad de Salamanca, Salamanca, 2011.

[5] L. Vygotsky, «El desarrollo de los procesos psicológicos superiores,» 1979.

[6] E. Turban, Decision Support and Expert Systems: Management Support Systems, Prentice Hall, 1995.

[7] E. M. Sánchez Vila y M. Lama Penín, «Técnicas de la Inteligencia Artificial Aplicadas a la Educación,» Inteligencia Artificial. Revista Iberoamericana, pp. 7-12, 2007.

[8] H. B. G. S. i. S. a. T. P. - J. J. K. S. i. E. Policy-, «Science, Technology, Engineering, and Mathematics (STEM) Education: A Primer.,» 1 Agosto 2012.

[9] D. Mar, «Robótica Aplicada a la Educación,» 2006.

[10] P. C. Lozano, «El Sistema Tutorial AgentGeom y su contribución a la mejora de las competencias de los alumnos en la resolución de problemas de matemáticas,» Universidad de Barcelona.

[11] V. Gaitán, «Gamificación: el aprendizaje divertido,» 2013. [En línea]. Available: <https://www.educativa.com/blog-articulos/gamificacion-el-aprendizaje-divertido/>.

[12] C. Barroso, «ENS-AI: Un sistema experto para la enseñanza,» de teoría de la educación, Universidad de La Laguna, 1994, pp. 237-251.

Estrategias innovadoras en los Procesos de Enseñanza y de Aprendizajes de Informática

Gustavo J. ASTUDILLO¹, Silvia BAST¹, Pedro A. WILLGING^{1,2}, Darío SEGOVIA¹, Leandro CASTRO¹ & Juan, DISTEL¹

¹ Departamento de Matemática/FCEyN/UNLPam

² CONICET

astudillo@exactas.unlpam.edu.ar, silviabast@exactas.unlpam.edu.ar,
pedro@exactas.unlpam.edu.ar, dariosegovia2000@gmail.com, leajcastro@gmail.com,
disteljm@gmail.com

Resumen

Desde el proyecto “Incorporación de Estrategias innovadoras en los Procesos de Enseñanza y de Aprendizajes de Informática”, nos proponemos investigar el impacto de diferentes estrategias tecno-pedagógicas sobre la enseñanza y el aprendizaje en algunos temas de informática. Dentro del gran espectro de posibilidades que brindan las actuales tecnologías para la enseñanza y el aprendizaje, orientaremos nuestra investigación sobre: la gamificación, los juegos serios y la robótica educativa. Desde lo pedagógico el foco estará en: el aprendizaje basado en problemas, el pensamiento computacional, los principios del buen aprendizaje y los estilos de aprendizaje.

Se desarrollarán estrategias innovadoras que definan secuencias de aprendizaje reutilizables que incluyan actividades, recursos y materiales para responder a la currícula de Informática. Una vez implementadas, se analizará cómo éstas influyen en el aprendizaje de los estudiantes y la motivación del docente.

Palabras clave: gamificación, juegos serios, robótica educativa, aprendizaje basado en problemas, estilos de aprendizaje

Contexto

En el marco de proyectos de investigación previos, completados a la fecha por este grupo

investigación GrIDIE (Facultad de Ciencias Exactas y Naturales - UNLPam), se han desarrollado secuencias de aprendizaje y materiales educativos; se han evaluado diferentes entornos virtuales de aprendizaje, y se ha analizado el impacto de distintas aplicaciones informáticas en el proceso educativo. También se investigaron distintos aspectos del paradigma de objetos de aprendizaje y de los Repositorios educativos/institucionales que los alojan (que derivó en la instalación del Repositorio Institucional, ReDi). Recientemente, el GrIDIE se enfocó también en el uso de gamificación y juegos serios para el diseño de ambientes educativos reales y virtuales.

La experiencia alcanzada en estos años de investigación nos ha permitido consolidar el grupo de trabajo y, en 2018, dos de las líneas de investigación se transformaron en sendos proyectos de investigación: “Aprendizaje de las ciencias con tecnologías educativas” y “Incorporación de Estrategias innovadoras en los Procesos de Enseñanza y de Aprendizajes de Informática” ambos con evaluación externa y aprobados por RCD 27/18.

Introducción

En este trabajo presentamos el proyecto de investigación “Incorporación de Estrategias innovadoras en los Procesos de Enseñanza y de Aprendizajes de Informática”. El mismo, se enfoca analizar el impacto de distintas

tecnologías y metodologías de enseñanza utilizadas y que tienen el potencial de mejorar el aprendizaje de distintos temas de las Ciencias Informáticas.

Como punto de partida para el proceso de investigación se pondrá el foco en la gamificación, los juegos serios y la robótica educativa; así como también en el aprendizaje basado en problemas, el pensamiento computacional, los principios del buen aprendizaje y los estilos de aprendizaje.

Estilos de Aprendizaje

Uno de los intereses actuales de la investigación en el ámbito educativo es indagar sobre los estilos de aprendizaje de los estudiantes. Los mismos, tiene una importancia sustancial para el diseño de la enseñanza y en la personalización de los aprendizajes. Felder & Silverman (1988) afirman que el aprendizaje de un estudiante se rige, por una parte, por habilidades propias que han sido adquiridas previamente, pero también por su estilo de aprendizaje y la compatibilidad de éste con el estilo de enseñanza del profesor. No existe un único modelo para estilos de aprendizaje (Zatarain & Barrón, 2011). Autores como Alonso, Gallego & Honey (1995) y Honey & Mumford (1986) definen cuatro categorías: activos, reflexivos, teóricos y pragmáticos. Mientras que Felder & Silverman (1988) utilizan: visual-verbal, sensitivo-intuitivo, secuencial-global, y activo-reflectivo. El modelo de la programación Neurolingüística (PNL) refiere tres estilos de aprendizaje y clasifica a los estudiantes en: visuales, auditivos y kinestésicos (Esquivel *et al.*, 2013).

Pensamiento Computacional

La enseñanza de la programación implica que aprender a analizar un problema, proponer posibles diseños de solución, seleccionar adecuado y probar/mejorar la solución. En síntesis, que el estudiante aprenda a resolver problemas y expresar la resolución en una secuencia eficiente de instrucciones.

Se trata de una competencia compleja, que implica abstracción, refinamiento, modularidad, entre otros. Compañ-Rosique *et al.* (2015) afirman “Para cualquier persona diseñar la solución a un problema requiere de un esfuerzo importante de abstracción, aún más si tiene que expresarla en forma de un algoritmo” (p.12). Para ello debe desarrollarse en los estudiantes el pensamiento computacional que, según los mismos autores, “se basa en resolver problemas haciendo uso de conceptos fundamentales de la informática” (p.1). Esto puede ser de suma utilidad, tanto para profesionales de la informática, como para quienes deben resolver un problema en sus vidas cotidianas.

Aprendizaje Basado en Problemas

El Aprendizaje Basado en Problemas (ABP) es un método de enseñanza y de aprendizaje centrado en el estudiante. El mismo, invierte la secuencia tradicional, presentando el problema a resolver como punto de partida para la adquisición de nuevos conocimientos. Según De la Torre *et al.* (2006), ABP incluye de manera implícita el desarrollo de habilidades, actitudes y valores benéficos para la mejora personal y profesional del estudiantes y promueve en ellos: gestión del conocimiento, la práctica reflexiva y la adaptación a los cambios.

Juegos Digitales y Educación

Prensky (2001) sostenía, casi dos décadas atrás, que nuestros estudiantes han cambiado radicalmente y ya no son las personas para las que el sistema educativo fue diseñado para enseñar.

Muchos de nuestros estudiantes, toman parte de su formación informal a través de videojuegos, simuladores, ambientes 3D, realidad virtual/aumentada/4D. Recursos que ya habían sido señalados como de potencial educativo por Informe Horizonte edición 2009 (Johnson, Levine & Smith, 2009).

En particular, "...los videojuegos representan un reto continuo para los usuarios que, además de observar y analizar el entorno, deben asimilar y retener información, realizar razonamientos inductivos y deductivos, construir y aplicar estrategias cognitivas de manera organizada y desarrollar determinadas habilidades psicomotrices (lateralidad, coordinación psicomotor, entre otras) para afrontar las situaciones problemáticas que se van sucediendo ante la pantalla. Aquí el jugador siempre se implica y se ve obligado a tomar decisiones y ejecutar acciones motoras continuamente..." (Marquès Graells, 2011).

Los juegos digitales pueden clasificarse según su objetivo en: (i) para entretenimiento, (como *Mario Brothers*) y (ii) juegos serios (como *Kokori*). Estos últimos, son juegos digitales diseñados para educar, entrenar o informar (Michael & Chen, 2005). Existen actualmente, una importante cantidad de este tipo de juegos orientados al aprendizaje de distintos conceptos de informática, particularmente la programación de computadoras (Astudillo, Bast & Willging, 2016).

También se ha tomado desde la dinámica de estos juegos, una estrategia para innovar en el diseño de la clase, la gamificación. La cual, hace referencia al "uso de elementos del diseño de juegos en contextos no lúdicos" (Deterding *et al.*, 2011, p.2). Como afirman Gallego *et al.* (2014) "gamificar es plantear un proceso de cualquier índole como si fuera un juego. Los participantes son jugadores y como tales son el centro del juego [...] y deben divertirse mientras consiguen los objetivos propios del proceso gamificado" (p. 2). Autores como McGonigal (2011) o Kapp (2012) coinciden en que los juegos digitales deben contar con las siguientes características: objetivos, reglas, desafíos o conflictos, competencia, colaboración y/o cooperación, retroalimentación y re-jugabilidad (*re-playability*). Las mismas deben ser tenidas en cuenta al llevar adelante un proceso de gamificación.

La aplicación de gamificación no se ciñe, únicamente, al ámbito educativo, sin embargo en este contexto pretende ir más allá de la

motivación, buscando promover los aprendizajes a partir de la resolución de problemas a través de la interacción con el ambiente y/o con otros jugadores.

Robótica educativa

Como se afirma en el Horizon Report - Edición Educación Superior 2016 "La noción trabajar y vivir entre los robots es cada vez menos futurista y más práctica que nunca" (Johnson *et al.*, 2016, p. 46). En este contexto, una tendencia que cobra fuerza es el uso de robots con fines educativos, la denominada robótica educativa. "La robótica se refiere al diseño y aplicación de robots, que son máquinas que realizan una serie de tareas automatizadas" (Johnson *et al.*, 2016, p. 46). Aplicado al contexto educativo, la robótica "forma parte de un enfoque pedagógico centrado en el alumno, que le permite construir objetos tangibles de su propio diseño y con sentido para él." (Vaillant, 2013, p. 38).

Para Bruner (1961) es necesaria la participación activa del estudiante en el proceso de aprendizaje, donde estudiante y docente cooperan en la resolución del problema. La robótica educativa entonces, convierte a la robótica en un medio para alcanzar ciertos aprendizajes.

Inicialmente, los robots cuentan con el potencial de facilitar el aprendizaje de un lenguaje de programación, propiciar la experimentación y estimular las competencias asociadas a la resolución de problemas, permite concretizar la relación existente entre software y hardware, y hacen posible además que el estudiante obtenga un resultado palpable de su producción, aportando entonces a la idea de "construccionismo", que afirma que el conocimiento se construye (Papert & Harel, 1991).

Líneas de Investigación, Desarrollo e Innovación

Este proyecto se desarrolla bajo la hipótesis de que es posible definir estrategias

innovadoras para la enseñanza y el aprendizaje de temas de informática, utilizando TIC, que impacten positivamente en la motivación de los estudiantes y en los diseños didácticos de los docentes.

La línea central de investigación se focaliza en el impacto que resulta de la integración de estrategias innovadoras nuevas o existentes, que involucren el uso de TIC, para la enseñanza y el aprendizaje de temas de informática.

Esto implicará investigar sobre la incorporación de juegos serios, los alcances de las robótica educativa, y de estrategias basadas en gamificación, así como también, cómo conjugar éstas con el aprendizaje basado en problemas y teniendo en cuenta los distintos estilos de aprendizaje, con el fin de desarrollar el pensamiento computacional y teniendo en cuenta los principios del buen aprendizaje.

Resultados y Objetivos

Los objetivos específicos del proyecto son:

- Explorar las potencialidades de diferentes tipos de juegos serios como mediadores del proceso de enseñanza y de aprendizaje.
- Releva, analizar y adaptar estrategias que permitan la gamificación de ambientes educativos .
- Releva, analizar y adaptar experiencias usen robótica educativa para la enseñanza y/o el aprendizaje de conceptos de informática.
- Medir el impacto de las estrategias seleccionadas en la motivación de docentes y estudiantes.
- Releva y evaluar entornos de desarrollo, de acceso libre, que permitan la programación desarrollar el pensamiento computacional.
- Releva y evaluar recursos y aplicaciones que permitan desarrollar actividades en ambientes simulados.
- Generar situaciones problemáticas, disponibles con licencia Creative Commons, desde el Repositorio ReDi.

- Analizar el impacto sobre el rol docente de la introducción de la metodología de ABP en algunas asignaturas del trayecto académico de los estudiantes.

Los resultados esperados de la consecución de este proyecto serán:

- Producción de estrategias innovadoras que permitan abordar distintos temas de informática a través de la inclusión de TIC.
- Definición de estrategias basadas en ABP que favorezcan el pensamiento computacional.
- Creación de recursos educativos abiertos.
- Incremento de la cantidad de materiales y recursos didácticos publicados en el Repositorio Institucional.
- Fortalecimiento del área de investigación en TIC aplicada en educación en nuestro medio.

Formación de Recursos Humanos

En este proyecto, un investigador formado y cuatro investigadores en formación, dos de los cuales alcanzaron el grado de Magíster durante 2016.

En el marco del nuevo Plan de estudios del Profesorado en Computación, los estudiantes, deben acreditar la participación en proyectos de investigación. Actualmente, el proyecto cuenta con una estudiante avanzada realizando tareas en ese contexto.

Referencias

- Alonso, C.M.; Gallego, D.L. y Honey, P. (1995).** *Los estilos de aprendizaje.* Procedimientos de diagnóstico y mejora. Bilbao: Ediciones Mensajero.
- Astudillo, G. J., Bast, S. G., & Willging, P. A. (2016).** Enfoque basado en gamificación para el aprendizaje de un lenguaje de programación. *Virtualidad, Educación y Ciencia*, 7(12), 125–142.
- Burbules, N. (2009).** Ubiquitous Learning: A New Sociocultural Context for Education. Texto de la conferencia, II Congreso Internacional Educación, Lenguaje y Sociedad. Universidad Nacional de La Pampa. General Pico.

- Bruner, J. S. (1961).** The act of discovery. *Harvard Educational Review*, 4, 21-32.
- Compañ-Rosique, P., Satorre-Cuerda, R., Llorens-Largo, F., Molina-Carmona, R., (2015).** Enseñando a programar: un camino directo para desarrollar el pensamiento computacional. *RED- Revista de Educación a Distancia*, 46(11). DOI:10.6018/red/46/11
- De la Torre Solarte, G.; Narváez Paredes, E.; Rosas Guevara, L.; Romo Roseo, A.; Fernández Sandoval, N.; Jiménez Chappotín, G.; Collazos Ordóñez, C.; Muñoz Rodríguez, D.; Erazo Santander, O.; Astaiza A, M.; Correa Correa Z (2006).** *Pensamiento Universitario - Propuesta Educativa*. Universidad Cooperativa de Colombia. Facultad de Educación. Segunda Edición. ISBN: 958-8205-42-5
- Deterding, S., Khaled, R., Nacke, L. E., & Dixon, D. (2011).** Gamification: Toward a definition. En *CHI 2011 Gamification Workshop Proceedings*. Vancouver, BC, Canada.
- Esquivel, P., Cantú, L., Cantú, M., Aguirre, D., & González, M. (2013).** Determinación de los estilos de aprendizaje de los estudiantes de una licenciatura del área química (pp. 1578-1592). Presentado en 1er. Congreso Internacional de Investigación Educativa, Monterrey, Nuevo León.
- Felder, R. M., & Silverman, L. K. (1988).** Learning and teaching styles in engineering education. *Engineering education*, 78(7), 674-681.
- Gallego, F., Molina, R., & Faraón, L. (2014).** Gamificar una propuesta docente. Diseñando experiencias positivas de aprendizaje. Presentado en XX Jornadas sobre la enseñanza universitaria de la informática, Oviedo, España.
- Honey, P. and Mumford, A. (1986a).** *The Manual of Learning Styles*, Peter Honey Associates.
- Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., y Hall, C. (2016).** *NMC Informe Horizon 2016 Edición Superior de Educación*. Austin, Texas: The New Media Consortium.
- Johnson, L., Levine, A., & Smith, R. (2009).** *Informe Horizon*. Austin, Texas: The New Media Consortium. Traducción al español de The 2009 Horizon Report.
- Marquès Graells, P. (2011, Agosto 7).** *LOS VIDEOJUEGOS. LOS VIDEOJUEGOS*. Página Web.
- McGonigal, J. (2011).** *Reality Is Broken: Why Games Make Us Better and How They Can Transform the World*. New York: The Penguin Press.
- Michael, D. R., & Chen, S. L. (2005).** *Serious Games: Games That Educate, Train, and Inform*. Muska & Lipman/Premier-Trade.
- Papert, S., & Harel, I. (1991).** Situating constructionism. *Constructionism*, 36(2), 1-11.
- Premsky, M. (2001).** Digital Natives, Digital Immigrants. *On the Horizon*, 5(9), MCB University Press.
- Vaillant, D. (2013).** Integración de TIC en los sistemas de formación docente inicial y continua para la Educación Básica en América Latina. Argentina: UNICEF Argentina.

Influencia del uso de la gamificación y las herramientas de evaluación continua en el proceso de enseñanza y aprendizaje

Horacio René Del Giorgio; Verónica Inés Aubin; Leonardo José Blautzik; Lucas Videla; Renata Guatelli; José Luis Cabrera; Carolina Sánchez; Alejandro Goitea

Departamento de Ingeniería e Investigaciones Tecnológicas

Universidad Nacional de La Matanza.

Florencio Varela 1903 - San Justo (CP 1754)

Tel: 4480-8952

hdelgiorgio@unlam.edu.ar; vaubin@unlam.edu.ar; leoblautzik@gmail.com; videla.lucas@gmail.com; renata.guatelli@gmail.com; icabrera@unlam.edu.ar; carolinafsanchez@gmail.com; agoitea@unlam.edu.ar

Resumen

Continuando con la idea de la cátedra de “Programación Avanzada” de la Universidad Nacional de La Matanza de incorporar metodologías activas para el proceso de enseñanza y aprendizaje, tales como la programación de a pares, el uso de PSP (Proceso Personal de Software), el trabajo colaborativo on-line, el aprendizaje basado en problemas y la clase invertida, se propone la idea de incorporar la evaluación continua y estudiar su efectividad para valorar la adquisición de las competencias esperadas en los estudiantes de Ingeniería Informática; en especial, la de trabajo en equipo.

El seguimiento continuo de los equipos de estudiantes que persiguen el logro de un objetivo conduce hacia una pronta detección de errores, tanto en los enfoques como en la realización, y ayuda a corregir rápidamente los esfuerzos mal canalizados. Esto permite, además, comprobar el correcto balance de la participación de todos los integrantes de un equipo, evitando injusticias y previniendo conflictos.

Siendo ésta la manera en que se realiza habitualmente el seguimiento de cualquier proyecto de software, la aplicación de estas estrategias didácticas específicas continúa con el propósito de la cátedra de incrementar las habilidades del estudiante en su desempeño académico y acercarlo de manera temprana a la actividad profesional.

Palabras clave: Competencias, educación en ingeniería, metodologías, trabajo en equipo, evaluación continua.

Contexto

Este Proyecto de Investigación tiene su fundamento en el PEICB (Proyecto Estratégico de Ingeniería para Ciencias Básicas) del DIIT (Departamento de Ingeniería e Investigaciones Tecnológicas) de la UNLaM (Universidad Nacional de La Matanza), en el marco del Programa de fortalecimiento de las carreras de Ingeniería de la Secretaría de Políticas Universitarias (2012-2016).

En el PEICB se propuso el desarrollo de estrategias pedagógicas superadoras que favorecieran procesos de aprendizaje significativos y la incorporación de enfoques basados en la formación por competencias tendientes a una mayor retención del alumnado. Para ello se recomendaron estrategias activas de enseñanza tales como aprendizaje basado en problemas, aprendizaje por proyectos, trabajo colaborativo, entre otras, como metodologías pedagógicas centrales en el desarrollo de las asignaturas.

El tiempo de realización de este proyecto es de dos años, desde comienzos de 2018 a finales del 2019

Introducción

El sistema educativo está atravesado por un conjunto de transformaciones sociales vinculadas al desarrollo de las TICs (Tecnologías de la Información y Comunicación). La adaptación de los procesos actuales a estos cambios implica que los modelos educativos a seguir deben someterse a una revisión, junto al papel de los participantes en dichos procesos, como así también los entornos donde se lleva a cabo el aprendizaje.

Según Pérez Lindo, para construir esta nueva visión compleja de los procesos educativos hay que comenzar por reconocer que han entrado en crisis las ideas de realidad y verdad, los paradigmas sobre la subjetividad, la visión de los entornos naturales y culturales y los principios que guiaban los métodos de enseñanza tradicionales [Pérez Lindo, 2009].

Por otra parte, el cambio de paradigma en la educación, de la forma tradicional a la centrada en el estudiante, implica la necesidad de actualizar también la forma de evaluar. Estas nuevas modalidades de trabajo se ven fortalecidas al implementar una evaluación continua que permite un seguimiento tanto individual como del equipo de trabajo.

El objetivo de incluir el trabajo colaborativo en la propuesta de la cátedra es crear situaciones en las cuales se generen interacciones productivas entre los estudiantes [Ronteltap & Eurelings, 2010]. Trabajar en equipos permite, a través de una función principalmente colaborativa, que los estudiantes adquieran destrezas interpersonales y cognitivas, así como habilidades que los capaciten para enfrentarse a distintas situaciones grupales a lo largo de su trayectoria académica y profesional.

Adoptar el sistema de evaluación continua como una estrategia de evaluación formativa orientada al proceso de aprendizaje en lugar de evaluaciones puntuales permite aumentar la motivación

del estudiante, valorar si el estudiante alcanzó no sólo los conocimientos esperados, sino también las competencias previamente definidas por el docente para una asignatura y tema en particular. Esta forma de trabajo también permite que la asimilación de conocimientos y el desarrollo de competencias por parte de los estudiantes se realice durante todo el proceso. Asimismo, aumenta en forma significativa el feedback entre el docente y los estudiantes.

Por otra parte, la incorporación de las herramientas GitHub, Overlaf, y Travis al proceso de evaluación continua, permiten mejorar la forma de evaluar y acompañar a los estudiantes. Gran parte de las actividades quedan registradas y con ello se torna más simple, ante un problema presentado por algún estudiante, trazar su actividad y detectar el origen del mismo.

La aplicación de herramientas de seguimiento que interactúen junto con GitHub o Travis, permite a su vez obtener mejoras sensibles en la relación con los estudiantes, ya que ahora poseen un feedback constante y en todo momento conocen en detalle su situación.

Durante el 2017, una primera implementación de lo arriba mencionado ha permitido un seguimiento exitoso en un Taller de Programación dictado en ambos cuatrimestres.

Dado que se puede supervisar el trabajo casi en tiempo real, se logra obtener información en cuanto a la participación de cada estudiante en el trabajo grupal, terminando con los desvíos típicos del trabajo en equipo, tales como “el que no delega” por miedo a que las cosas se hagan mal y “el vago”, o “háganlo ustedes, yo me adoso”, y logrando que situaciones como éstas y otras similares queden de alguna manera registradas.

Las mejoras en las que se está trabajando, que de hecho son la base del presente Proyecto de Investigación, son la aplicación de la evaluación continua a todas

las tareas evaluables del curso. Se pretende incorporar mecánicas de juego a las actividades educativas con el fin de aumentar la motivación, y además incorporar herramientas de seguimiento que contemplen el uso de PSP y permitan centralizar las mediciones que los estudiantes tomen de su trabajo. De esta manera será posible computar y calcular los promedios, desvíos y percentiles de las muestras para realizar un análisis más pormenorizado de cada una de las actividades evaluables y su impacto en los estudiantes, ya sea de manera individual y/o grupal.

Líneas de Investigación, Desarrollo e Innovación

El presente proyecto de investigación da continuidad a diferentes proyectos y trabajos desarrollados en el DIIT de la UNLaM, en el campo de la enseñanza en la educación superior. Algunos de ellos son los siguientes:

C135. Reorganización disciplinar y didáctica de Matemática Discreta para las carreras de Ingeniería. (2011-2012).

C137. Uso de nuevas métricas orientadas a las competencias en la gestión curricular. (2012-2013).

C161. Análisis y propuestas de estrategias didácticas innovadoras en el dictado de Asignaturas relacionadas con las TICs. (2014-2015).

C167. Mejoras en el proceso de enseñanza y aprendizaje de Programación utilizando metodologías propias de la industria del software como caso particular de las metodologías activas. (2014-2015).

C191. Fortalecimiento de competencias transversales de trabajo en equipo. (2016-2017).

Para el desarrollo del presente proyecto, se ha conformado un grupo interdisciplinario, integrado por Especialistas en Educación e Ingenieros

Informáticos y Electrónicos, que aportarán un amplio conocimiento sobre distintos enfoques del proceso de enseñanza y aprendizaje en asignaturas relacionadas con la Programación, lo cual permitirá la contribución de estrategias didácticas innovadoras específicas para esta disciplina [Camilloni, 2007]. Dichos profesionales conforman actualmente la planta de docentes-investigadores de la UNLaM.

Resultados y Objetivos

Los modelos de enseñanza y aprendizaje orientados al desarrollo de competencias conducen a la necesidad de evaluar el grado de adquisición de las mismas. El objeto de evaluación ha dejado de ser exclusivamente la cantidad de conocimientos adquiridos, debiendo incluir habilidades y actitudes desarrolladas. Por otro lado, la evaluación de competencias debe acompañar el proceso de aprendizaje y no exclusivamente evaluar un resultado final y ello lleva a la necesidad de sistematizar el seguimiento de las experiencias de aprendizaje de los estudiantes (Durán Aponte & Durán García, 2014).

La evaluación continua se adopta como una estrategia de evaluación formativa más orientada al proceso de aprendizaje que a una valoración cuantitativa de objetivos. Para que la evaluación continua pueda llevarse a cabo a lo largo del curso, deben proponerse actividades de carácter evaluable por parte de los docentes.

En el caso de este Proyecto de Investigación, la orientación será hacia la asignatura “Programación Avanzada”, lo cual implica, tal como ya se mencionó anteriormente, una focalización mayor aún, ya que se estará trabajando con elementos de Didáctica de la Especialidad [Camilloni, 2007].

Con respecto a los resultados esperados, se investigará sobre diversas aproximaciones para mejorar los resultados en la enseñanza de la programación a partir

de la aplicación de técnicas de trabajo en equipo. Para ello, durante los dos años de duración del Proyecto (2018-2019), se han asumido los siguientes compromisos, que se pueden resumir del siguiente modo:

- Se realizará una Investigación documental pormenorizada de los antecedentes y del estado del arte en técnicas de desarrollo de software, en especial las de trabajo colaborativo. Asimismo, también se indagará sobre herramientas libres que se encuentran disponibles para desarrollar software trabajando en equipo y que permitan hacer seguimiento y control de versiones.
- Se realizará un relevamiento sobre las cátedras de la Carrera de Ingeniería en Informática relacionadas con la Asignatura “Programación Avanzada” en donde se trabaja en equipo y este trabajo debe ser evaluado. También se recopilará todo el material didáctico utilizado en el dictado y evaluación de la asignatura con el objetivo de su futura adaptación a las modificaciones metodológicas que se espera proponer. De ser posible, también se realizarán relevamientos en asignaturas similares en otras universidades.
- Basado en todo lo anterior, se diseñarán herramientas de Software para que, interactuando con las ya existentes, se puedan implementar las mejoras derivadas que surjan de los pasos anteriores para poder así optimizar la gestión de la evolución de los estudiantes, tanto en su rendimiento individual como grupal.
- A continuación, se desarrollará un caso práctico experimental en el que se puedan contrastar las diferencias obtenidas con la

implementación de las rutinas de Software diseñadas previamente; todo esto a través de entrevistas y encuestas a estudiantes con el fin de poder evaluar, a posteriori, los resultados del nuevo ambiente de seguimiento que se ha creado.

- Finalmente, y basado en todo lo anterior, se elaborará un Documento de Recomendaciones de Buenas Prácticas para la evaluación del Trabajo en Equipo en asignaturas relacionadas con la “Programación Avanzada”.

Formación de Recursos Humanos

El grupo de estudiantes y profesores del DIIT involucrados en el proyecto resultarán directamente beneficiados con estos desarrollos en el aspecto académico y curricular.

Estos conocimientos se podrán volcar en cursos específicos de Capacitación Docente dirigidos a los docentes de las cátedras del Departamento de Ingeniería de la UNLAM, a fin de promover la aplicación de metodologías de aprendizaje colaborativo, metodologías ágiles y promover la discusión de nuevas técnicas de evaluación.

También está previsto el desarrollo de una Tesis de Maestría en Informática a cargo de uno de los integrantes del Grupo.

Bibliografía

- Bannan-Ritland, B., Dabbagh, N., & Murphy, K. (2000). Learning Object Systems as Constructivist Learning Environments: Related Assumptions, Theories, and Applications. Obtenido de https://www.researchgate.net/publication/265996930_Learning_Object_Systems_as_Constructivist_Learning_Environments_Related_Assumptions_Theories_and_Applications

- Bernad, J. (2000). Modelo cognitivo de evaluación educativa, Escala de estrategias de aprendizaje contextualizado. Madrid: Narcea.
- Cabrera, F. (2003). Evaluación de la formación. Madrid: Síntesis.
- Camilloni, A. (2007). El Saber Didáctico. Buenos Aires: Paidós.
- Durán Aponte, E., & Durán García, M. (2014). Competencias sociales y las prácticas profesionales. Vivencias y demandas para la formación universitaria actual. Obtenido de <https://www.tandfonline.com/doi/abs/10.1174/113564012799740777>
- Ilgen, D., Hollenbeck, J., Johnson, M., & Jundt, D. (2005). Teams in Organizations: From Input-Process-Output Models to IMO Models. Obtenido de [http://faculty.washington.edu/mdj3/Ilgen,%20Hollenbeck,%20Johnson,%20&%20Jundt%20\(2005\).pdf](http://faculty.washington.edu/mdj3/Ilgen,%20Hollenbeck,%20Johnson,%20&%20Jundt%20(2005).pdf)
- Mathieu, J., Maynard, M., Rapp, T., & Gilson, L. (2008). Team Effectiveness 1997-2007: A Review of Recent Advancements and a Glimpse Into the Future. Obtenido de https://www.researchgate.net/publication/220041231_Team_Effectiveness_1997-2007_A_Review_of_Recent_Advancements_and_a_Glimpse_Into_the_Future
- Pérez Lindo, A. (2009). Para qué educamos hoy. Buenos Aires: Biblos.
- Rico, R., Alcover de la Hera, C., & Taberner, C. (2010). Efectividad de los Equipos de Trabajo, una Revisión de la Última Década de Investigación (1999-2009). Obtenido de http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1576-59622010000100004
- Ronteltap, F., & Eurelings, A. (2010). Activity and Interaction of Students in an Electronic Learning Environment for Problem-Based Learning. Obtenido de <https://www.tandfonline.com/doi/abs/10.1080/01587910220123955?journalCode=cdie20>

INTEGRACIÓN DE SERVICIOS EN APLICACIÓN MÓVIL PARA APRENDIZAJE UBICUO

**Mariana BRACHETTA, Julio MONETTI, Bianca SOZZI, Mercedes MUÑOZ
& Oscar LEÓN**
Ingeniería en Sistemas de Información/FR.Mendoza -UTN
{mbrachetta,julio.monetti,biancasozzi,mercedesmp1305,oleon111}@gmail.com

RESUMEN

En el trabajo se describen aspectos centrales considerados en el diseño, desarrollo e implementación de una aplicación móvil que integra geolocalización, servicios en segundo plano, tecnología de realidad aumentada (RA) y servicios de computación en la nube para lograr una experiencia concreta y medible de aprendizaje ubicuo.

Palabras clave: *Aprendizaje Ubicuo, Geolocalización, Realidad Aumentada, Mobile Cloud Computing Android, Entornos de Aprendizaje.*

CONTEXTO

Un prerequisite al proponer una experiencia de aprendizaje ubicuo [1] es ofrecer al alumno actividades educativas situadas, es decir, disponibles en el sitio y momento donde se encuentra. Se conciben así las actividades en contexto como factor clave de todo aprendizaje [2].

El teléfono móvil integrado a los servicios en red, representa en este sentido una herramienta potencialmente adecuada para proponer a los estudiantes este tipo de actividades.

El presente trabajo se enmarca en el proyecto de investigación *PID UTN4741 (Desarrollo de un entorno basado en Cloud Computing para Aprendizaje Ubicuo)*. La integración de esta aplicación en un entorno de aprendizaje basado en *Cloud Computing*, permitirá luego a un equipo de docentes del área de Ciencias Básicas, proponer actividades de aprendizaje ubicuo y evaluar en qué medida los alumnos

adquieren conocimiento a través de los elementos que encuentran en el entorno de su vida cotidiana, es decir, llevar el aprendizaje fuera de las aulas.

Se resumen, en este trabajo los resultados alcanzados en investigación y desarrollo preliminar del proyecto, durante el cual se profundizaron los conocimientos teóricos necesarios para la implementación de la mencionada aplicación.

1. INTRODUCCION

En el presente trabajo se exponen resultados referentes a la tarea de formulación y resultados parciales obtenidos en el marco del proyecto mencionado. Se diseña y codifica una aplicación para el sistema Android [3] que hace uso de la geolocalización del alumno para presentar conceptos a través de realidad aumentada dependiendo la ubicación de este. La aplicación utiliza los servicios proporcionados por la pila de software de la arquitectura Android. Los servicios más relevantes utilizados, los cuales también se encuentran bajo estudio y análisis, son: Administrador de Ubicaciones, Administrador de Notificaciones, OpenGL|ES, Controlador de Wifi, Controlador de la Cámara, Controlador de Pantalla, Controlador de teclado. Para completar la funcionalidad deseada se utiliza:

- Para Georeferenciación: **GoogleMaps API** [5], desarrollado en el apartado 2.1, y **Servicios en Segundo Plano**, que se puede encontrar en el apartado 2.2.

- Para Realidad Aumentada: **Librería Vuforia** [6], ampliado en el apartado 2.4. y **motor gráfico Unity** [7] explicado en el apartado 2.5.

2. LINEA DE INVESTIGACIÓN Y DESARROLLO

2.1. Manejo de GPS en un entorno Android

Es requisito funcional de la aplicación poder identificar cuando un usuario (alumno), se encuentra en un determinado lugar geográfico en el que existen objetos de aprendizaje a su disposición.

La geolocalización del alumno se aborda con el uso de las librerías *Google Maps*.

Las funcionalidades de ubicación, creación de mapas, agregación de marcadores, etc., son aportadas en entornos Android por *Google Play Services*. Dado que ni los emuladores de ejecución Android disponibles, tales como *GenyMotion*, ni el entorno de desarrollo utilizado (*Android Studio*), disponen por defecto de estos servicios, los mismos deben instalarse en el SDK (Kit de desarrollo Android) de manera adicional.

El sistema Android provee dos mecanismos alternativos para la geolocalización: `FINE_LOCATION` y `COARSE_LOCATION`, para georeferenciar con distintos niveles de precisión la ubicación del dispositivo. El uso de estos mecanismos se configura en el manifiesto de la aplicación (archivo *AndroidManifest.xml*). Por defecto se utiliza el mecanismo de `FINE_LOCATION` (alta precisión con uso de GPS), puesto que se torna necesaria la ubicación precisa del alumno en un momento dado. En caso de no disponerse el servicio en tiempo de ejecución el sistema accede automáticamente a la modalidad `COARSE_LOCATION` (menor precisión con uso de datos de la conexión Wi-Fi o de la red de telefonía celular). Para acceder al GPS, se requieren permisos de usuario provistos en tiempo de instalación de la aplicación [8].

2.2. Servicios en segundo plano

Cuando el alumno ingresa en un ámbito geográfico donde existen objetos de aprendizaje accesibles (polígono de interés), es preciso notificar de tal evento, dado que es un

requerimiento de la aplicación llamar la atención del estudiante, para proponerle actividades de aprendizaje diseñadas para un sitio y contexto determinado. Para implementar el servicio de notificación se utilizan los servicios en segundo plano de Android. Esta característica, propia del sistema operativo, permite la ejecución ininterrumpida de un proceso de la aplicación que reacciona ante la ocurrencia de un evento [9].

Existen varias opciones para implementar tareas de segundo plano en Android:

- *AsyncTask*
- *Thread*
- *Handler*
- *Service*

Todas las opciones antes mencionadas están materializadas a través de clases Java provistas en el SDK de Android. Se realizaron pruebas comparando las diferentes alternativas, optándose por el uso de *Service* para dar solución a los requisitos funcionales descriptos. Un *Service* constituye un componente de la aplicación, que permite la ejecución de tareas de larga duración en segundo plano, sin la presencia de una interface gráfica activa. Con su uso se logra que el servicio esté permanentemente en funcionamiento, esperando que se produzca el evento. Existen dos tipos de *Service*:

- Servicio iniciado
- Servicio de enlace

La diferencia entre ambos radica en el tipo de vinculación que existe entre el servicio y el resto de los componentes de aplicación o incluso su conexión con aplicaciones de servidor. En función de los requisitos funcionales definidos, se optó por un servicio iniciado. En particular, el servicio permanece de forma indefinida *escuchando* los cambios de ubicación y reacciona ante el ingreso del dispositivo al polígono de interés, y notificando finalmente al alumno de tal situación.

Para permitir al servicio *escuchar* los cambios de ubicación en tiempo real se implementa la interface Java *LocationListener* [10]. Para generar el mensaje de notificación al alumno

se utilizan las clases de Android *Builder* y *NotificationManager*, junto a otras utilizadas para configurar el comportamiento de la notificación. Mediante la notificación se le presentan al alumno contenidos educativos disponibles en el contexto en que se encuentra, preparados especialmente por el equipo docente para promover actividades de aprendizaje oportunas y situadas.

Se evaluaron dos opciones alternativas para determinar cuando el usuario ingresa al polígono de interés:

Primera Opción: Definir la zona de interés como una circunferencia, ubicando un punto central mediante coordenadas y haciendo comparaciones dentro de un radio. Esta opción resulta conveniente porque es sencilla de implementar y requiere menor capacidad de almacenamiento de datos. Su desventaja radica en la menor precisión con que se determina el polígono de interés, el cual queda necesariamente inscrito dentro de la circunferencia.

Segunda Opción: Definir la zona de interés como un polígono uniendo un conjunto de puntos. La ventaja de esta opción es su mayor precisión para definir el área donde ubicar los objetos de aprendizaje (Por ejemplo: el ámbito de una habitación, el *buffete* o el campo de juego universitario). Su desventaja radica en la necesidad de guardar mayor cantidad de coordenadas y requerir mayor capacidad de cómputo para determinar el ingreso del usuario en la zona de interés.

Dado que tanto la disponibilidad de memoria como las capacidades de procesamiento pueden ser limitadas en un dispositivo móvil, se decidió que para los objetivos de la experiencia la primera opción resulta más conveniente. Así, cuando el alumno ingresa en el área de interés definida y visualiza la notificación correspondiente, la aplicación activa una segunda fase que hace uso de la tecnología de RA [4] para ofrecer al alumno actividades que aprovechan aspectos del ámbito real donde se encuentra; y lo enriquecen con contenidos especialmente

preparados para favorecer situaciones de aprendizaje significativo.

2.3. Realidad Aumentada

El concepto de RA está relacionado con cómo las personas perciben la realidad, y de qué forma la tecnología puede enriquecer dicha experiencia. La RA necesita cuatro elementos básicos:

- un **dispositivo** para capturar imágenes (cámara)
- un **elemento sobre el que proyectar las imágenes** (celular, computadora, consola de videojuegos)
- un **software de procesamiento**, capaz de interpretar la información que provee el usuario a través de su dispositivo y generar la información virtual correspondiente
- un **elemento “activador de realidad aumentada”**. Este concepto se refiere a la imagen ante la cual el sistema debe reaccionar (*Target*)

Se analizó el estado del arte con respecto a diferentes librerías utilizadas en proyectos actuales.

2.4. Librería Vuforia

Para el desarrollo de la aplicación se utilizó la librería Vuforia, que es una plataforma para la creación de aplicaciones de RA con soporte multiplataforma para los sistemas operativos *Mobile*.

La librería ofrece múltiples tipos de reconocimiento, tales como identificación de texto, imágenes, *multi-targets* y *cylinder-targets*. El presente trabajo aborda sólo lo relativo a reconocimiento de imágenes. Los *Image Targets* son imágenes planas sobre las cuales se presentan los objetos. Las imágenes que se utilizan como objetivo deben guardarse en una base de datos a la que Vuforia accede para poder comparar la imagen capturada por la cámara con la información almacenada y reaccionar según corresponda.

Las imágenes deben estar en formato JPG o PNG, en color o blanco y negro. Podría definirse que un *target* es adecuado o no a los

efectos de su reconocimiento, de acuerdo al número de *features* o puntos de reconocimiento que tenga la imagen. Por ejemplo, ver figura 1:

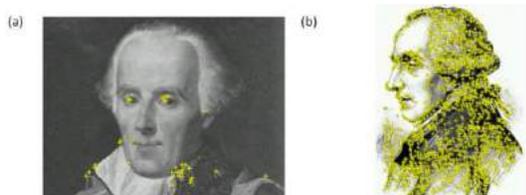


Figura 1 – Image Target. (a) Mal Target (b) Buen Target

Los *features* son los puntos en color amarillo. Como se puede observar, la primera imagen (Figura 1.a) sólo contiene unos pocos, por lo que podría ser calificada como un mal *target*, lo que quiere decir que será más difícil su reconocimiento con el dispositivo.

La librería Vuforia se complementa con el motor gráfico Unity. Este último permite generar contenido dinámico y proyectarlo sobre la imagen original.

2.5. Integración de Unity y Vuforia

Unity es un motor de videojuego multiplataforma. Permite el desarrollo de juegos para múltiples dispositivos tales como computadoras, consolas de videojuego, smartphones, smart TV, dando soporte además a los distintos sistemas operativos que cada uno de ellos puede utilizar. Unity 2017.2 provee el motor de Vuforia integrado. En caso de utilizar otra versión, el paquete puede importarse desde Unity, descargándose desde el portal de desarrollador de la página oficial de Vuforia. Una vez importado el paquete, se pueden utilizar los elementos de la librería, incluyendo *targets* predefinidos que pueden utilizarse.

Toda aplicación de Vuforia requiere una clave de licencia. Dicha licencia es gratis durante el proceso de desarrollo. Una vez que la aplicación se pone en producción es necesario actualizar la licencia a una versión paga, la cual debe registrarse en la configuración de Vuforia dentro de Unity.

Dentro de la configuración del *ImageTarget* debe establecerse la base de datos que se utilizará y qué imagen dentro de ella será el objetivo en ese caso.

El comportamiento de los objetos dentro de Unity se maneja a través de *scripts* en lenguaje C#. El ejemplo analizado hasta el momento fue la inclusión de texto sobre el objeto en el momento en que se detecta el *target*. Para ello, se debe modificar el script *Default Trackable Event Handler* y agregar el comportamiento deseado dentro de los métodos Java *OnTrackingFound()* y *OnTrackingLost()*.

Una vez realizada la configuración correspondiente para el comportamiento deseado el resultado obtenido al ejecutar la aplicación es el siguiente:

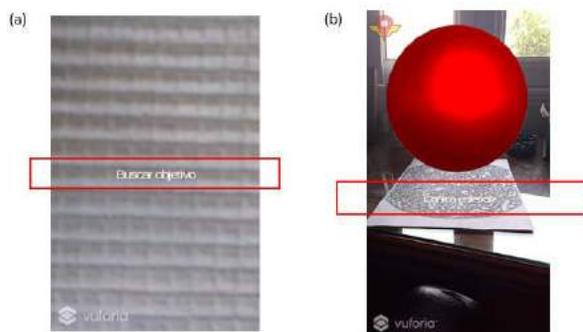


Figura 3 – Resultado de la ejecución. (a) Antes de detectar el objetivo –Texto: “Buscar Objetivo” (b) Objetivo detectado – Texto: “Cónica esférica”.

3. EMPAQUETAMIENTO Y DISTRIBUCIÓN DE LA APLICACIÓN

El desarrollo, codificado como un proyecto en el entorno Android Studio, se exporta y distribuye como un único fichero en formato ejecutable para sistema Android (*apk*).

Una vez obtenido el fichero *.apk*, el mismo debe ser firmado digitalmente para su correcta instalación y uso en dispositivos Android.

El paquete completo puede ser publicado en el repositorio *Google PlayStore* o distribuido de manera personalizada a cada uno de los usuarios. En una primera instancia del proyecto la aplicación será publicada y distribuida desde un servidor propio de la Universidad a fin de que el cuerpo docente

tenga mayor control sobre quienes acceden al mismo.

4. RESULTADOS OBTENIDOS

Se ha desarrollado un prototipo que incorpora los conceptos de geolocalización y RA para la presentación de contenido dinámicamente. De forma complementaria, el prototipo incorpora la conexión a un servidor para identificar al alumno y manejar las sesiones de usuario.

El prototipo desarrollado, permite que el equipo docente -no especializado en el área de informática- pueda visualizar el alcance y las posibilidades que el uso de estas tecnologías, adecuadamente integradas, proporciona la posibilidad de implementar experiencias concretas de aprendizaje ubicuo.

5. FORMACION DE RECURSOS

El equipo de trabajo está conformado por profesores y alumnos avanzados de la carrera *Ingeniería en Sistemas de Información* de la Universidad Tecnológica Nacional, Facultad Regional Mendoza. Los becarios se están formando en temas técnicos con la posibilidad de incorporarse a la carrera de investigador, e integrar nuevas actividades de investigación relacionadas con el presente trabajo.

6. BIBLIOGRAFIA

[1] Burbules, Nicholas. El aprendizaje ubicuo y el futuro de la enseñanza. *Encounters on Education*. Vol. 13, 2012, 3 – 14. University of Illinois, Urbana-Champaign, Estados Unidos. ISSN 1925-8992.

[2] Sagástegui, Diana. Una apuesta por la cultura: el aprendizaje situado. *Sinéctica, Revista Electrónica de Educación* N° 24. ITESO Universidad Jesuita de Guadalajara, Jalisco, México. (2004). Disponible en <http://www.sinectica.iteso.mx/index.php/SINECTICA/article/view/282>.

[3] Cruz Zapata, Belen. *Android Studio Essentials*. Packt Publishing. 2015. Birmingham, UK. ISBN 978-1-78439-720-3.

[4] Gazcón, Nicolás F, Larregui, Juan I, Castro, Silvia M. *La Realidad Aumentada*

como complemento motivacional. *Libros Aumentados y Reconstrucción 3D*. TE&ET N°17 [pág. 7-15]. RedUNCI-UNLP. 2016. ISSN 1850-9959.

[5] GoogleMaps API. Disponible en <https://developers.google.com/maps/documentation/android-api/current-place-tutorial>.

[6] Página Oficial de Vuforia. Disponible en <https://library.vuforia.com/>.

[7] Página Oficial de Unity. Disponible en <https://unity3d.com/es>.

[8] Solicitud de permisos. Disponible en <https://developer.android.com/training/permissions/requesting.html?hl=es-419>.

[9] Servicios en segundo Plano de Android. Disponible en <https://developer.android.com/guide/components/services.html?hl=es-419>.

[10] LocationListener. Disponible en <https://developer.android.com/reference/android/location/LocationListener.html>.

[11] Zapata-Ros, Miguel. Calidad en entornos ubicuos de aprendizaje. *RED. Revista de Educación a Distancia*. Número 31. Universidad de Alcalá. España.

[12] Linowes J. *Unity Virtual Reality*. Packt Publishing. Birmingham, UK. 2015. ISBN 978-1-78398-855-6.

SISTEMA DE PRODUCCIÓN POLIMEDIA FA.C.E.N.

Marcelo Omar Sosa¹, Raúl Marcelo Vega², Eugenia Cecilia Sosa Bruchmann¹

¹Departamento Computación/ ² Departamento Química/Facultad de Ciencias Exactas y Naturales/Universidad Nacional de Catamarca

Av. Belgrano N° 300 - Planta alta - C.P: 4700 - San Fernando del Valle de Catamarca
Teléfono: 0383- 4425610 /4420900

sosamodl@hotmail.com, vegarm@gmail.com, sosab_ec@hotmail.com

Resumen

Las nuevas tecnologías se incorporan paulatinamente en cada actividad humana en forma casi imperceptible, se han convertido en la actualidad en elementos imprescindibles de la vida cotidiana. Con la masificación de ellas, las actividades humanas se han visto modificadas en casi todos sus aspectos, desde el vocabulario de las personas comunes hasta la forma de enseñar y aprender. Es por ello, que la educación como una actividad humana global y en particular la de nivel superior representa un campo propicio para la incorporación de la tecnología como herramienta de mejora continua. Muchos de los aportes que realizan las tecnologías al proceso educativo van evolucionando y retroalimentándose de las experiencias que se desarrollan en diferentes partes del mundo. El acceso a internet permite compartir estas experiencias facilitando la replicación de aquellas que se consideran exitosas. En particular el desarrollo del sistema de producción polimedia, el cual se basa en el desarrollo de materiales educativos de calidad, representa una innovación en el proceso educativo. Este sistema se pretende replicar en la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de Catamarca. Las características de los materiales educativos desarrollados en este sistema redundan en ventajas para la transmisión de conocimiento sobre un tema específico, permitiendo

aprovechar al máximo la motivación que genera en los alumnos.

Palabras clave: materiales educativos, sistema de producción polimedia, educación superior

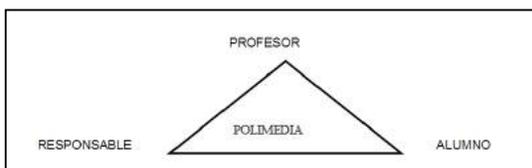
Contexto

El presente trabajo de investigación se desarrolla dentro de las líneas prioritarias de la Facultad de Ciencias Exactas y Naturales (FA.C.E.N) de la Universidad Nacional de Catamarca. Se enmarca como proyecto de investigación el que se encuentra en proceso de evaluación en el sistema de acreditaciones en la Secretaría del Ciencia y Técnica de la misma Universidad. Los alumnos a los que estará destinado inicialmente pertenecen a distintas carreras informáticas de la FA.C.E.N. relacionadas con la enseñanza de la programación. Los docentes investigadores están a cargo de asignaturas de las carreras del área informática que se dictan en el campus central en la Capital de la provincia de Catamarca como así también en subsedes del interior Catamarqueño (Departamento Ancasti) y en la provincia de Tucumán (San Miguel de Tucumán).

Introducción

El sistema de producción polimedia es una creación de la Universidad Nacional de

Valencia – España. Su objetivo principal es permitir la creación de materiales educativos de calidad que servirán para el apoyo a la docencia presencial y virtual. Abarca desde la preparación del material docente, pasando por la grabación hasta la distribución y puesta a disposición para alumnos y usuarios en general. Se denomina sistema porque va más allá de un conjunto de dispositivos para la creación de materiales educativos, está compuesto por numerosos servicios para los docentes y para la organización eficiente de los recursos técnicos y físicos necesarios para su funcionamiento. Existen numerosas universidades que desarrollaron sus sistemas polimedias y que vienen generando materiales propios relacionados con las disciplinas que se imparten en cada una de ellas. Estas experiencias y materiales pueden ser consultados y sirven como punto de partida para el objetivo de este trabajo. Como toda institución de nivel superior, la FA.C.E.N. conlleva como principal objetivo el relacionado con el rendimiento académico de sus alumnos, por ello se pretende trabajar en una mejora continua del proceso educativo. En este sentido, las acciones se orientan a la incorporación de la tecnología en forma innovadora dentro de dicho proceso. En el sistema de producción polimedia pueden separarse las funciones en dos aspectos bien diferenciados, por un lado el concerniente al equipamiento tecnológico y espacio físico y por el otro, el referido al diseño de materiales y organización general del funcionamiento del sistema. Además participan en él el personal que hace funcionar el sistema que está compuesto por los operadores técnicos que aseguran el correcto funcionamiento y asesoran a los docentes en la grabación de los materiales. Y los docentes o expertos que se responsabilizan de las tareas académicas.



Participantes del sistema polimedia, figura extraída de:
<https://ddd.uab.cat/pub/dim/16993748n18/16993748n18a5.pdf>

El sistema de producción polimedia necesita más detalladamente tener en cuenta los siguientes aspectos.

-Desde el aspecto del espacio físico o estudio polimedia, éste debe contar con el equipamiento y dispositivos electrónicos que permita la creación de los materiales educativos. Estos elementos se enumeran a continuación:

- Una computadora para la grabación y edición de videos.
- Pantalla auxiliar para que el docente pueda ver los contenidos.
- Cámara de video para la grabación en alta definición.
- Pantalla fondo blanco para facilitar la edición y superposición de imágenes en el video.
- Iluminación adecuada a la grabación en alta definición de los videos.
- Micrófono personal para la captación con fidelidad del relato del docente.
- Atril para posicionar los dispositivos a utilizar en el manejo de las presentaciones por parte del docente.
- Software para la edición del video.
- Televisión auxiliar para la visión lateral de los contenidos expuestos por parte del docente durante la exposición.

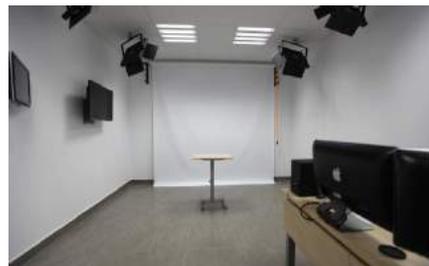


Imagen de la sala polimedia. Extraída de la página de Universidad Católica de Murcia

-Desde el aspecto humano, debe contarse con el técnico de grabación que estará a cargo del manejo del equipamiento durante la exposición del docente. Además deberá asesorar al docente en aspectos técnicos como tiempos estipulados, tipo de vestimenta, posicionamiento frente a la pantalla,

movimientos adecuados, tiempos y énfasis en el relato entre muchos otros imprevistos que surgen en la grabación.

El docente debe preparar el material a desarrollar el que puede contener videos, presentaciones, planillas de cálculo e incluso tiene la posibilidad de realizar dibujos o escritura en tiempo real. Para facilitar el desarrollo de la exposición el docente debe diseñar un libreto en donde organizará los contenidos en los cuales pueden realizarse anotaciones y comentarios que solamente podrán ser vistos por él y que le servirán de guía durante la grabación.



Imagen de la exposición del docente. Extraída de la página de Universidad Católica de Murcia

-Desde el aspecto del diseño de los contenidos se debe establecer características comunes que deberán cumplir los materiales para ser aceptados en la biblioteca polimedia.

Estas condiciones se presentan a continuación para facilitar su comprensión:

- Establece una plantilla oficial que contenga las características pretendidas por la Facultad.
- Planificar la exposición según los siguientes pasos Apertura, cuerpo o desarrollo del tema y conclusión.
- Establecer el tamaño de las presentaciones, por lo general el formato 16:9.
- Aconsejar la vestimenta del expositor, en color y tipo de modo que no afecten la edición o el contenido del video.
- Tener en cuenta que el discurso del expositor puede ser presentado en las pantallas de forma que actúe como prompter.

- Aconsejar la duración del video, por lo general no sobrepasar los 10 minutos.
- Utilizar en el contenido imágenes propias o de uso libre.

Metodología

Se pretende con esta investigación diseñar y poner en funcionamiento el espacio para la producción de materiales educativos polimedia de la FA.C.E.N.. Inicialmente se trabajará con carreras de informática para generalizar luego a las distintas carreras que pertenecen a la Facultad. Se desarrollarán cursos y capacitaciones tanto en diseño de materiales como para el manejo del equipamiento. La meta principal consiste en que cada carrera cuente con el personal capacitado para la generación material educativo de calidad que estará a disposición para los alumnos en el website de la FA.C.E.N..

Conclusión

Se espera replicar las características del sistema de producción polimedia desarrollado por numerosas universidades Europeas y en otras partes del mundo, estableciéndose las bases necesarias para los primeros desarrollos polimedia. Se espera lograr un trabajo incremental iniciando con los contenidos básicos de programación de modo que permitan validar el procedimiento y los resultados obtenidos. Con los objetivos de mejorar el rendimiento académico de los alumnos y generar materiales educativos de calidad basados en la innovación en el proceso educativo en el nivel superior.

Líneas de investigación y desarrollo

El presente trabajo se desarrolla dentro de las actividades previstas en el proyecto de investigación denominado "Innovación tecnológica: Sistema polimedia FA.C.E.N.". Para el desarrollo de los materiales multimedia de calidad se utilizan inicialmente los temas más relevantes de cada unidad de las asignaturas con contenidos de programación

de las diferentes carreras de la Facultad. Para el análisis de los resultados del proyecto se registrarán como datos iniciales el desempeño de los alumnos del grupo de control en actividades seleccionadas, los que se compararan con los datos finales del mismo grupo cuando utiliza el sistema polimedia.

El presente tema se viene desarrollando mediante el estudio continuo con la presentación de diferentes trabajos en reuniones científicas en donde se muestran las primeras acciones realizadas y prototipos desarrollados puestos en funcionamiento. Además de analizar los posibles resultados que se esperan de la investigación. En general se encuentra dentro de los temas seleccionados para el diseño de proyectos de investigación.

Resultados y Objetivos

El resultado esperado para esta investigación es la creación y puesta en funcionamiento del sistema de producción polimedia FA.C.E.N.

Tiene como objetivos principales:

- Conformar un grupo crítico de personal capacitado para el manejo adecuado del equipamiento que permita asesorar a los docentes durante la grabación de los materiales educativos.
- Diseñar el espacio físico para la ubicación del sistema de producción polimedia de la FA.C.E.N.
- Seleccionar el equipamiento necesario como así también el software que permita la grabación y edición de los materiales educativos.
- Establecer las pautas necesarias para el diseño de los materiales y el libreto para que sirva de guía a los docentes.
- Establecer grupos de control para la obtención de datos que servirán para la cuantificar los resultados obtenidos.

- Establecer los criterios de selección de los contenidos teóricos que servirán de bases para los materiales educativos de calidad.

Formación de Recursos Humanos

Los autores del trabajo se encuentran en la etapa de desarrollo de sus tesis de posgrado en carreras relacionadas con el tema de investigación, como la Maestría en Ciencias de la Computación en donde el Mgter. Sosa, Marcelo Omar Diógenes desarrolla actualmente la tesis denominada: “*Aportes de data mining a la mejora del proceso educativo con blended learning: formalización experimentaciones*” que se realiza con la dirección del Dr. Chesñevir Carlos Iván perteneciente a la Universidad Nacional del Sur (U.N.S). Además el investigador se encuentra en la etapa de planificación de su tesis doctoral en el área de minería de datos en el Doctorado en Ciencias dictado en la Facultad de Ciencias Exactas y Naturales (FA.C.E.N.) en convenio con la Universidad Nacional del Sur (U.N.S).

La Docente Investigadora Especialista Sosa Bruchmann, Eugenia Cecilia continúa desarrollo su tesis en la carrera Maestría en Ingeniería en Software de la Universidad Nacional de San Luis. Los docentes investigadores desarrollan actividades de dirección de tesis de la carrera de Licenciatura en Tecnología Educativa de los siguientes alumnos: Varela Marino del Valle cuya tesis se denomina “Análisis del impacto de un aula virtual en el proceso de enseñanza y aprendizaje en la escuela de educación técnica N° 7 “ José Alsina Alcobert”, y Campagnale Rubén Darío cuya tesis se denomina: “Enseñanza mediada por las TIC: Una práctica posible en el nivel secundario de la escuela N° 54 Anexo I, La Carrera Departamento Fray Mamerto Esquiú- Provincia de Catamarca”. El Docente investigador Vega Raúl Marcelo desarrolla sus actividades docentes y de investigación en el departamento de Química de la FA.C.E.N.

Además desarrollan las siguientes

actividades:

- Dirección de proyectos de investigación acreditados por la secretaría de Ciencia y técnica de la Universidad Nacional de Catamarca.
- Integración de equipo de investigación de centro de investigación en Estadística de la Facultad de Ciencias Exactas y Naturales del U.N.Ca.
- Producción de artículos científicos para su presentación en congresos locales, nacionales e internacionales.
- Participación de los integrantes en cursos de actualización y posgrado en el área de estudio.
- Integrantes de la revista de ciencias de la Facultad de Ciencias Agrarias de la U.N.Ca.
- La actualización y capacitación permanente de los investigadores en talleres o workshop relacionadas con el tema del trabajo.
- La participación de los investigadores como consultores en proyecto afines que se desarrollan en la Facultad de Ciencias exactas y Naturales en distintas áreas.
- Examinadores de trabajos finales en las diferentes carreras que se dictan en la FA.C.E.N. de las U.N.Ca.
- Dirección de tesis y tesinas finales de las carreras de computación, informática y Licenciatura en tecnología educativa.
- La planificación de seminarios para docentes en temas relacionados con la investigación y resultados obtenidos en la investigación.
- Dictado de cursos para docentes investigadores sobre temas desarrollados en proyectos e investigaciones sobre el tema del trabajo.

BIBLIOGRAFÍA Y TRABAJOS CONSULTADOS

- A. Dasso, et. al. Teaching Programming. En Proceedings de las Primeras Jornadas de Educación en Informática y TICS en Argentina, (2005), Bahía Blanca, Argentina,

pp.183-187.

- A. Ferreira Szpiniak, G. Rojo, Cambios metodológicos didácticos y evaluación del impacto de los mismos en un curso introductorio a los conceptos de algorítmica y programación. En Proceedings de las Primeras Jornadas de Educación en Informática y TICS en Argentina, (2005), Bahía Blanca, Argentina, pp.210-216.
- Cabero Almenara J., María del Carmen Llorente Cejudo (2007). "Propuestas de colaboración en educación a distancia y tecnologías para el aprendizaje". Eduotec. Revista Electrónica de Tecnología Educativa. Núm. 23.
- Diaz Barriga F: (2011). "La innovación en la enseñanza soportada en TIC. Una mirada al futuro desde las condiciones actuales". VII Foro Latinoamericano de Educación / Experiencias y aplicaciones en el aula. Aprender y enseñar con nuevas tecnologías.
- Prendes Espinosa, M. P., Martínez. (2006)."Actividades individuales versus actividades colaborativas", en E-actividades: un referente básico para la formación en Internet, ISBN 84-665-4768-1. 183-202.
- Becker, K.; Marquardt, C.G.; Ruiz, D.D. A. (2004).Pre-Processing Tool for Web Usage Mining in the Distance Education Domain. pp. 78-87.
- Pahl, C. (2004). Data Mining Technology for the Evaluation of Learning Content Interaction. International Journal on E-Learning IJEL.
- Eliana Esther Gallardo Echenique Universidad Continental (Perú).

Aplicación de “píldoras de conocimiento” a través del sistema polimedia en el ámbito educativo

- Guardiola, P. (2012). Los contenidos audiovisuales como mejora en la calidad de la enseñanza universitaria: el caso de iTunes U. In D. Cobos, A. Jaén, E. López, A. Martín, & L. Molina (Eds.), I Congreso Virtual Internacional sobre Innovación Pedagógica y Praxis Educativa INNOVAGOGÍA 2012 (pp. 1206–1216). Sevilla: Universidad Pablo de Olavide. Disponible en: <http://www.upo.es/ocs/index.php/innovagogia2012/Iinnovagogia2012/paper/download/123/125>.
- Gutiérrez Castillo, J. J., & Cabero-Almenara, J. (2015). De la resentación al polimedia. una experiencia en ciencias de la educación. Revista Educação, Cultura E Sociedade, 5(2), 28–40.
- Infante Moro, A., Santos Fernández, N., Muñoz Ronchel, C., & Pérez Medina, L. (2010). Aplicación del polimedia en el ámbito educativo. DIM: Didáctica, Innovación Y Multimedia, 1-7(18).
- Maceiras, R., Cancela, Á., & Goyanes, V. (2010). Aplicación de Nuevas Tecnologías en la Docencia Universitaria. Formación Universitaria, 3(1), 21–26. doi:10.4067/S0718-50062010000100004

SITIOS Y PAGINAS WEB

- www.ucam.edu/servicios/polimedia/qu-e-es-polimedia. *accedido en marzo de 2018*.
- www.cted.uclm.es/servicios/polimedia. *accedido en marzo de 2018*.
- polimedia.uab.cat/. *accedido en marzo de 2018*.
-

HERRAMIENTAS DE INTERPRETACIÓN GRÁFICA PARA CODIFICACIÓN DE DIAGRAMAS DE MODELADO DE SISTEMAS INTERPRETABLES POR UN DISMINUIDO VISUAL

Hernán Amatriain, Sebastian Martins, Santiago Bianco, Federico Ribeiro, Nicolás Pérez

Grupo Investigación en Sistemas de Información
Laboratorio de Investigación y Desarrollo en Ingeniería del Software y en Sistemas de Inteligencia Artificial
Departamento de Desarrollo Productivo y Tecnológico, Universidad Nacional de Lanús
29 de Septiembre 3901 (1826) Remedios de Escalada, Lanús, Argentina. Tel +54 11 5533 5600 Ext. 5194
hamatriain@gmail.com

RESUMEN

La actual Ley Nacional de Educación Superior hace responsable al Estado del acceso a la educación universitaria a todo aquel que quiera hacerlo y cuente con la formación y capacidad requerida, y garantizar la accesibilidad al medio físico, servicios de interpretación y los apoyos técnicos necesarios y suficientes, para las personas con discapacidad [Ley 25573, 2002]. También confiere a las instituciones universitarias la autonomía académica e institucional para formular y desarrollar planes de estudio, de investigación científica y de extensión y servicios a la comunidad incluyendo la enseñanza de la ética profesional y la formación y capacitación sobre la problemática de la discapacidad. Sin embargo los desarrollos en algunas áreas de menor demanda aún son escasos.

Existen pocas herramientas para que los disminuidos visuales trabajen con diagramas de modelado de sistemas. Las principales [Planttext; Plantuml; PSeInt], incluso son incompletas en cuanto por un lado pueden generar gráficos a partir de un código, pero no pueden realizar el trabajo inverso; y por otro lado, no abarcan todos los diagramas de modelado de sistemas, dejando afuera, por ejemplo, los Diagramas Entidad Relación (DER) que se utilizan para la modelización de datos (para la posterior creación de la Base de Datos).

El presente proyecto busca desarrollar un conjunto completo de herramientas para que un no vidente pueda interactuar con un grupo de estudio o docente (en un ambiente de estudio o evaluación) o un grupo de desarrollo (en un ambiente laboral) a través de estos diagramas de modelado de sistemas.

Palabras clave: Modelado de sistemas para disminuidos visuales. Lenguaje interpretación gráfica. Herramienta diseño de software. No vidente.

CONTEXTO

El presente proyecto busca cubrir la vacancia existente en un conjunto de herramientas que den soporte a los no videntes para comunicar ideas a través de diagramas de sistemas, es decir, un conjunto de herramientas de software que haga de intérprete para que una persona con discapacidad visual pueda intercambiar opiniones utilizando gráficos de modelado de sistemas como ser los Diagramas de Flujo, Diagramas Entidad-Relación, Lenguaje Unificado de Modelado, Grafos, Arboles, Redes de Petri, etc. Al no existir en la actualidad un conjunto completo de estas herramientas, no puede llevarse a cabo la inclusión completa de personas con discapacidad visual (no vidente) en la Educación Superior. El objetivo principal de este proyecto es desarrollar este conjunto de herramientas para hacer de intérprete para que una persona ciega pueda interactuar a través de diagramas de sistemas.

INTRODUCCIÓN

El 30 de abril de 2002 se publica en el Boletín Oficial la Ley N° 25573, modificatoria de la Ley 24521, Ley de Educación Superior (Boletín Oficial N° 29888). Entre las modificaciones que introduce esta ley, está la que hace referencia a la “responsabilidad indelegable” del Estado de la prestación del servicio de educación superior de carácter público, que debe garantizar la accesibilidad al medio físico, servicios de

interpretación y los apoyos técnicos necesarios y suficientes para las personas con discapacidad [Ley 25573, 2002]. Otras modificaciones que incorpora son las de asegurar el derecho de estudiantes con discapacidad de acceder a los servicios de interpretación y los apoyos técnicos necesarios y suficientes durante las evaluaciones, y la de “formar y capacitar científicos, profesionales, docentes y técnicos, capaces de actuar con solidez profesional, responsabilidad, espíritu crítico y reflexivo, mentalidad creadora, sentido ético y sensibilidad social, atendiendo a las demandas individuales, en particular de las personas con discapacidad”.

En la Universidad Nacional de Lanús, se viene haciendo un trabajo de interpretación y apoyo técnico a personas con discapacidad desde el Programa de Inclusión Universitaria para Personas con Discapacidad de la Dirección de Bienestar Universitario dependiente de la Secretaría de Cooperación y Servicio Público. Esta iniciativa se fundamenta en la necesidad de abordar una problemática social más amplia, que remite a las oportunidades educativas que poseen aquellos sectores de la población históricamente excluidos del Sistema de Educación Superior, partiendo de una nueva manera de concebir, diseñar e implementar políticas universitarias, asentada en el reconocimiento, respeto y afirmación de las diferencias inherentes a la población estudiantil, congruente con la política de democratización del conocimiento que lleva adelante la Universidad Nacional de Lanús. Hasta hace poco tiempo la discapacidad era pensada como una cuestión que concernía exclusivamente a las personas afectadas, quienes con asistencia y protección podían rehabilitarse a fin de adaptarse al medio y lograr la integración social. Centrado en los aspectos médicos, en lo patológico, este modelo médico-céntrico de la discapacidad, perdía de vista los derechos de este grupo heterogéneo, como personas y como ciudadanos, ignorando que lo que está en juego es la calidad de vida y la construcción de una sociedad realmente inclusiva, capaz de enriquecerse de la diversidad.

Sin embargo, se ha presentado un caso que pone en evidencia la falta de madurez que se tiene en

ciertas áreas específicas con discapacidades puntuales. Particularmente se presentó una prueba difícil al equipo interdisciplinario cuando el primer estudiante no vidente de la carrera de Licenciatura en Sistemas de la universidad, tuvo que aprender y ser evaluado en algunas técnicas y metodologías de modelado de sistemas que son esencialmente gráficas como los Diagramas de Flujo Nassi-Shneiderman, Diagramas Entidad Relación (SER), Lenguaje Unificado de Modelado (UNL), que son solo algunas de las herramientas gráficas que forman parte curricular de la mencionada carrera (acreditada por Resolución CONEAU 1089/12). Efectivamente, no existen herramientas ni procedimientos especiales para evaluar a un no vidente en estas áreas específicas.

Se han investigado herramientas de apoyo para estas situaciones especiales, encontrando que las soluciones existentes [PSeInt, Planttext, PlantUML] son parciales e insuficientes. PSeInt es una herramienta para asistir a un estudiante en sus primeros pasos en programación. Mediante un simple e intuitivo pseudo lenguaje en español (complementado con un editor de diagramas de flujo), le permite centrar su atención en los conceptos fundamentales de la algoritmia computacional, minimizando las dificultades propias de un lenguaje y proporcionando un entorno de trabajo con numerosas ayudas y recursos didácticos. Si bien es una herramienta no diseñada para ciegos, puede utilizarse para la comprensión inicial de algoritmos para no videntes haciendo una introducción parcial a los diagramas Nassi-Shneiderman. PlantText es una herramienta en línea que genera rápidamente imágenes desde el texto. Principalmente, se utiliza para generar diagramas UML (Unified Modeling Language). PlantUML es un proyecto Open Source (código abierto) que permite escribir rápidamente: Diagramas de Secuencia, Diagramas de Casos de uso, Diagramas de Clases, Diagramas de Actividades, Diagramas de Componentes, Diagramas de Estados, Diagramas de Objetos. El inconveniente con las herramientas mencionadas es que son incompletas en cuanto por un lado pueden generar gráficos a partir de un código, pero no pueden realizar el trabajo

inverso, haciendo imposible que un ciego interprete un diagrama; y por otro lado, no abarcan todos los diagramas de modelado de sistemas, dejando afuera, por ejemplo, los Diagramas Entidad Relación (DER) que se utilizan para la modelización de datos (para la posterior creación de la Base de Datos).

En este contexto es que el presente proyecto busca diseñar un proceso de enseñanza y evaluación de no videntes para que puedan comunicar ideas a través de diagramas de modelado de sistemas por medio de un intérprete y desarrollar las herramientas software necesarias para darle al estudiante mayor independencia y al docente más agilidad para la enseñanza y evaluación.

PREGUNTAS PROBLEMA, HIPÓTESIS Y OBJETIVOS GENERALES Y ESPECÍFICOS

Pregunta de Estudio:

¿Es posible desarrollar un conjunto de herramientas de software que sirva de intérprete para que un disminuido visual pueda interactuar a través de diagramas de modelado de sistemas con otros individuos, sean ciegos o no?

Hipótesis:

Existe un emprendimiento donde se logra que a partir de un lenguaje desarrollado a tal fin, un ciego pueda realizar gráficos UML. Si bien es una herramienta muy útil, la misma es incompleta en cuanto por un lado pueden generar gráficos a partir de un código, pero no pueden realizar el trabajo inverso; y por otro lado, no abarcan todos los diagramas de modelado de sistemas, dejando afuera los DER que se utilizan para la modelización de datos (para la posterior creación de la Base de Datos). Por ello se supone posible desarrollar un conjunto completo de herramientas para que un no vidente pueda interactuar con un grupo individuos (sean ciegos o no) a través de estos diagramas de modelado de sistemas.

Objetivos:

1. Desarrollar un lenguaje de interpretación gráfica que pueda codificar los diagramas de modelado de sistemas existentes más utilizados de forma tal que pueda ser

interpretado por un no vidente, comprendiendo un gráfico así especificado o que pueda escribir el código que genera un diagrama específico.

2. Desarrollar un proceso de evaluación basado en el lenguaje de interpretación gráfica para que los docentes cuenten con un protocolo de actuación a la hora de evaluar a un disminuido visual en el área de utilización de diagramas de modelado de sistemas.
3. Desarrollar un conjunto de herramientas que utilicen el lenguaje de interpretación gráfica para codificar diagramas de modelado de sistemas y que, a la inversa, genere el código de un modelo específico graficado según la técnica que se esté utilizando.

METODOLOGÍA DE TRABAJO

Para construir el conocimiento asociado al presente proyecto de investigación, se seguirá un enfoque de investigación clásico [Riveros y Rosas, 1985; Creswell, 2002] con énfasis en la producción de tecnologías [Sábato y Mackenzie, 1982]; identificando métodos, materiales y abordaje metodológico necesarios para desarrollar el proyecto:

a. Métodos:

- Revisiones Sistemáticas: Las revisiones sistemáticas [Argimón, 2004] de artículos científicos siguen un método explícito para resumir la información sobre determinado tema o problema. Se diferencia de las revisiones narrativas en que provienen de una pregunta estructurada y de un protocolo previamente realizado.
- Prototipado Evolutivo Experimental (Método de la Ingeniería): El prototipado evolutivo experimental [Basili, 1993] consiste en desarrollar una solución inicial para un determinado problema, generando su refinamiento de manera evolutiva por prueba de aplicación de dicha solución a casos de estudio (problemáticas) de complejidad creciente. El proceso de refinamiento

concluye al estabilizarse el prototipo en evolución.

b. **Materiales:** Para el desarrollo de los formalismos y procesos propuestos se utilizarán:

- Formalismos de modelado conceptual usuales en la Ingeniería de Software [Rumbaugh et al., 1999; Jacobson et al., 2013].
- Modelos de Proceso usuales en Ingeniería de Software [IEEE, 1997; ANSI/IEEE, 2007; Oktaba et al., 2007].
- Herramientas Computación Gráfica [Edward, 2000; Foley et al 1995].
- Formalismos para expresar Gramáticas y Lenguajes [Beardon et al, 1991; Grishman, 1991; Moreno Sandoval, 1998; Fernández Fernández, 1995; Shieber, 1998; Chomsky, 1956, 1957, 1965, 1986, 2003 y 2008].

Normativas vigentes sobre las responsabilidades del Estado [Ley Educación Superior, 2002; Ley Accesibilidad Web, 2010].

RESULTADOS ESPERADOS

Como resultados de esta investigación se espera:

- Disponer de un lenguaje de interpretación gráfico que pueda codificar los diagramas de modelado de sistemas existentes más utilizados de forma tal que pueda ser interpretado por un no vidente.
- Disponer de un proceso de evaluación basado en el lenguaje de interpretación gráfica para que los docentes cuenten con un protocolo de actuación a la hora de evaluar a un disminuido visual en el área de utilización de diagramas de modelado de sistemas.
- Disponer de un conjunto de herramientas que utilicen el lenguaje de interpretación gráfica para codificar diagramas de modelado de sistemas y que, a la inversa, genere el código de un modelo específico graficado.

FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo se encuentra formado por un investigador formado y cuatro investigadores en formación. En su marco se desarrollan tres Tesis de Licenciatura en Sistemas.

FINANCIAMIENTO

Las investigaciones que se proponen en esta comunicación cuentan con financiamiento como Proyecto de Investigación 80020170400005LA de la Secretaría de Ciencia y Técnica de la Universidad Nacional de Lanús (Argentina).

REFERENCIAS

- ANSI/IEEE, (2007). Draft IEEE Standard for software and system test documentation. ANSI/IEEE Std P829-2007.
- Argimón J. (2004). Métodos de Investigación Clínica y Epidemiológica. Elsevier España, S.A. ISBN 9788481747096.
- Basili, V. (1993). The Experimental Paradigm in Software Engineering. En Experimental Software Engineering Issues: Critical Assessment and Future Directions (Ed. Rombach, H., Basili, V., Selby, R.). Lecture Notes in Computer Science, Vol. 706. ISBN 978-3-540-57092-9.
- Beardon, C; Lumsden, D; Holmes, G (1991). Natural Language and Computational Linguistics. An introduction. Ellis Horwood. Berge, N. & Berge, Z. (1998). Integration of Disabled Students into Regular Classrooms in the United States and in Victoria, Australia. The Exceptional Child, 35 (2), 107-117.
- Chomsky, Noam. (1956). Three models for the description of language. IRE Transactions PGIT, 2:113-124.
- Chomsky, Noam. (1957). Syntactic Structures. Mouton and co., N.Y. publishers.
- Chomsky, Noam. (1965). Aspects of the Theory of Syntax. The MIT Press, Cambridge, Massachusetts.
- Chomsky, Noam. (1986). Knowledge of Language: Its Nature, Origins and Use. Praeger Publishers, N.Y., USA.

- Chomsky, Noam. (2003). *Sobre la naturaleza y el lenguaje*. Cambridge University Press.
- Chomsky, Noam. (2008). On phases. En *Foundational Issues in Linguistic Theory*. The MIT Press, Cambridge, Massachusetts, páginas 133–166.
- Creswell, J. (2002). “Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research”. Prentice Hall. ISBN 10: 01-3613- 550-1.
- Edward, A. (2000) *Interactive Computer Graphics: a Top-Down Approach With OpenGL* 2nd edition, AddisonWesley, 2000.
- Fernández Fernández, G.; Sáez Vacas, F. (1995). *Fundamentos de informática: [lógica, autómatas, algoritmos y lenguajes]*. Editorial Anaya Multimedia.
- Foley, J.D.; van Dam, A.; Feiner, S.K.; Hughes, J.F. (1995). *Computer Graphics: Principles and Practice*, 2nd edition, AddisonWesley 1995
- Grishman, R. (1991). *Introducción a la Lingüística Computacional*. Visor.
- IEEE, (1997). *IEEE Standard for Developing Software Life Cycle Processes*. IEEE Std 1074-1997 (Revisión of IEEE Std 1074-1995; Replaces IEEE Std 1074.1-1995)
- Ley de Educación Superior (2002). *Ley Nacional Nro. 25573 (modifica Ley 24521)*. Publicada en el Boletín Oficial del 30-abr-2002, Número: 29888, Página: 1
- Ley de Accesibilidad web (2010). *Ley 26.653 de accesibilidad de la información en las páginas web*. Resolución 69/2011.
- Moreno Sandoval, A. (1998). *Lingüística Computacional. Síntesis*.
- Oktaba, H., Garcia, F., Piattini, M., Ruiz, F., Pino, F., Alquicira, C. (2007). *Software Process Improvement: The Competisoft Project*. IEEE Computer, 40(10): 21-28. ISSN 0018-9162.
- Planttext: para diagramas UML. Sitio oficial disponible en <https://www.planttext.com>.
- Plantuml: para diagramas UML. Sitio oficial disponible en <http://plantuml.com>.
- PSeInt: para diagramas Nassi-Schneiderman. Sitio oficial disponible en <http://pseint.sourceforge.net>
- Riveros, H. y Rosas, L. (1985). *El Método Científico Aplicado a las Ciencias Experimentales*. Editorial Trillas. México. ISBN 96-8243-893-4.
- Rumbaugh, J., Jacobson, I., Booch, G. (1999). *The Unified Modeling Language, Reference Manual*. Addison Wesley, ISBN-10: 02-0130-998-X.
- Sabato J, Mackenzie M. (1982). *La Producción de Tecnología: Autónoma o Transnacional*. Instituto Latinoamericano de Estudios Transnacionales - Technology & Engineering. ISBN 9789684293489.
- Shieber, S.M. (1989). *Introducción a los Formalismos Gramaticales de Unificación*. Teide.

PROCESOS DE DISEÑO DE COMPONENTES PARA ESPACIOS VIRTUALES DE TRABAJO ORIENTADOS A LA EDUCACIÓN DE PERSONAS CON DISCAPACIDADES

Darío Rodríguez, Roberto García, Rodolfo Priano, Norberto Charczuk, Santiago Bianco, Hernán Merlino

Laboratorio de Investigación y Desarrollo en Ingeniería de Explotación de Información
 Grupo de Ingeniería de Espacios Virtuales y Grupo Investigación en Sistemas de Información
 Departamento de Desarrollo Productivo y Tecnológico, Universidad Nacional de Lanús
 29 de Septiembre 3901 (1826) Remedios de Escalada, Lanús, Argentina. Tel +54 11 5533 5600 Ext. 5194
 tut77tec@gmail.com

RESUMEN

La evolución de las comunicaciones basadas en la tecnología de Internet es la base de la tendencia de desarrollo de los Espacios Virtuales de Trabajo (EVT), los cuales necesitan de un abordaje ingenieril que defina distintos procesos constructivos y de gestión (planificación y seguimiento de control). La normativa vigente respecto a la educación de nivel superior en Argentina y de accesibilidad Web, hacen que estos procesos deban incluir las herramientas, ayudas técnicas, y metodologías constructivas que tengan en cuenta la inclusión de personas con discapacidad a la hora de evaluar la usabilidad de estos EVT. El presente proyecto busca desarrollar un proceso de diseño de dispositivos educativos y evaluación basado en Espacios Virtuales de Trabajo y que atiendan de manera integral las necesidades de personas con distintas discapacidades cumpliendo con la normativa vigente.

Palabras clave: Espacios Virtuales de Trabajo, Ingeniería de Espacios Virtuales de Trabajo, Ingeniería Informática Aplicada en Educación, Accesibilidad en Sistemas Educativos.

CONTEXTO

En este proyecto se busca definir dos procesos y realizar dos implementaciones: [a] Proceso de Diseño de Dispositivos Educativos Centrados en EVTs y orientado a personas con discapacidad, [b] Diseño de Dispositivos Educativos Centrados en EVTs y orientado a personas con discapacidad en el ámbito de la UNLa, [c] Proceso de Evaluación de EVTs Orientados a Educación y orientado a personas con discapacidad, y [d] Implementar el Proceso de Evaluación de EVTs Orientados a Educación y orientado a personas con discapacidad en el ámbito de la UNLa.

El Grupo de Ingeniería de Espacios Virtuales de Trabajo de la Universidad Nacional de Lanús, desde el año 2010 viene desarrollando distintas herramientas conceptuales para cubrir la vacancia identificada. A la fecha se ha concentrado en formalismos de análisis y diseño conceptual de EVTs.

INTRODUCCIÓN

Los Espacios Virtuales de Trabajo (EVT) están destinadas a facilitar la mediación en el interior de equipos cuyos miembros no están físicamente contiguos [Gibson y Cohen, 2003], y tienen que desarrollar un objeto conceptual (por ejemplo: investigación, desarrollo de proyectos software, artículos técnicos, informes, documentación de diseño de edificios, planes de negocio, planes de inversión corporativos, entre otros). El EVT debe satisfacer el requisito de mantener y documentar las diferentes versiones del objeto conceptual que está siendo desarrollado por el equipo de trabajo de colaboración; dejando constancia de la evolución del acuerdo entre los miembros del grupo de trabajo desde las especificaciones iniciales del objeto conceptual hasta su etapa final de desarrollo.

En [Mikropoulos y Natsis, 2011], tras analizar una década de investigación con trabajo de campo en el área de los EVT de uso educativo, se señalan entre otras, la tendencia de sistematizar la integración de nuevos artefactos a los EVT (como los objetos de aprendizaje), así como dotarlos de un marco metodológico para su uso en el ámbito educativo y la evaluación de su usabilidad por la comunidad educativa [Redfern y Naughton, 2002].

La inclusión de personas con discapacidad en la educación superior pasó a ser una responsabilidad del Estado [Ley Educación Superior, 2002; Ley Accesibilidad Web, 2010], por lo que cualquier marco metodológico para el uso de los EVT en el ámbito educativo debe tener en cuenta estas características especiales, para que doten a los

objetos de aprendizaje con estos requerimientos esenciales e inclusivos.

CONVENCIONES, PREGUNTAS PROBLEMA, HIPÓTESIS Y OBJETIVOS GENERALES Y ESPECÍFICOS

Convenciones

Término: "Objeto de Aprendizaje"

Un Objeto de Aprendizaje es "una colección de contenidos, ejercicios, y evaluaciones que son combinados sobre la base de un objetivo de aprendizaje simple" [Wayne Hodgins, 2002]. En el contexto de los EVT orientados a Educación, un Objeto de Aprendizaje es un conjunto de recursos digitales, autocontenido y reutilizable, con un propósito educativo y constituido por al menos tres componentes internos: contenidos, actividades de aprendizaje y elementos de contextualización [Vessey y Conger, 1994]. El Objeto de Aprendizaje debe tener una estructura de información externa (metadatos) que facilite su almacenamiento, identificación y recuperación [Tejada et al., 2001]. Brito [2010] sostiene que a partir de la masiva y vertiginosa irrupción de las Tecnologías de la Información y Comunicación, los recursos diseñados deben ser accesibles, portables y mantenibles, para dar respuestas asequibles a las necesidades actuales. Focalizando sobre estos aspectos, la filosofía de Objetos de Aprendizaje (OA) en conjunto a la de Patrones de Diseño de OA, se erigen como pilares fundamentales en el desarrollo de materiales tecno-pedagógico-comunicacionales libres; proponiendo un enfoque integral que posibilita emplear todo el potencial de las TIC y, por lo tanto, incrementar y extender los beneficios de aplicación en la Educación a Distancia.

Término: "Dispositivos Educativos Centrados en EVTs"

Las tecnologías digitales han reconfigurado significativamente la Identidad, la Intimidad y la Imaginación durante las últimas décadas" [Gardner y Davies, 2014], por lo que los cambios culturales que se han producido en los estudiantes a partir de la interacción con esta revolución digital plantea nuevos interrogantes. En tal sentido, habrá que repensar cuáles serán los mejores dispositivos y escenarios para el aprendizaje y cómo ampliar el alcance de las aulas para lograr las competencias esperadas en los estudiante ya que, según Gardner y Davies [2014], "los jóvenes de ahora no solo crecen rodeados de aplicaciones sino que además han llegado a entender el mundo como un conjunto de

aplicaciones, a ver sus vidas como un conjunto de aplicaciones ordenadas". Del mismo modo se debe encontrar cómo potenciar nuestra enseñanza a través de los recursos tecnológicos disponibles en esta sociedad en red y cómo organizar y mediar pedagógica y eficientemente los contenidos de nuestros cursos. El a docentes debe incorporar competencias para crear o bien transformar sus cursos en espacios educativos mediados por tecnologías Web.

Las dificultades que se presentan para adaptar los cursos educativos a un escenario con TIC son complejas ya que "tanto las posibilidades que ofrecen las TIC para la enseñanza y el aprendizaje, como las normas, sugerencias y propuestas de uso pedagógico y didáctico de las mismas, son siempre irremediamente reinterpretadas y reconstruidas por los usuarios, profesores y alumnos, de acuerdo con los marcos culturales en los que se desenvuelven" [Monereo y Coll, 2008].

Desde el propio campo informático un equipo de la Universidad Jaime I – España propone una metodología para seleccionar tecnologías Web 2.0 para la docencia [Grangel et al., 2012], y en la Argentina la Dirección de Educación a Distancia de la Universidad Nacional de La Plata formula pautas de trabajo para propuestas educativas mediadas por tecnologías digitales [González et al., 2012].

Término: "Evaluación de EVT Educativos"

Centrando el tema en los espacios web, Marqués Graells [1999] analiza las principales funciones que pueden realizar en el ámbito educativo, presenta diversas tipologías y establece un modelo para la identificación y evaluación de estos espacios considerando diversos criterios de calidad.

Ferreira Szpiniak y Sanz presentan diferentes modelos de evaluación de EVEAs comparándolos, y revisan el concepto de usabilidad y de heurísticas de usabilidad [Ferreira Szpiniak y Sanz, 2007]; por otra parte, proponen modelo de evaluación de Entornos virtuales de enseñanza y aprendizaje [Ferreira Szpiniak y Sanz, 2012].

Término: "Accesibilidad Web"

La accesibilidad Web significa que personas con algún tipo de discapacidad van a poder hacer uso de la Web. La accesibilidad Web engloba muchos tipos de discapacidades, incluyendo problemas visuales, auditivos, físicos, cognitivos, neurológicos y del habla [W3C]. La W3C creó la iniciativa WAI -Web Accessibility Initiative- la cual se encarga específicamente de dictar normas de Accesibilidad como la WCAG 1.0 vigente desde 1999 y la WCAG 2.0, desde 2008, brindando los recursos necesarios para aplicarlas [W3C; WAI].

Preguntas Problema:

Con base en los principios de la Ingeniería de Espacios Virtuales de Trabajo (IEVT) y la normativa vigente respecto a la inclusión en la educación superior, ¿se pueden plantear los procesos de Diseño de Dispositivos Educativos Centrados en EVT's y de Evaluación de EVT's Orientados a Educación orientado a personas con discapacidad física?

Hipótesis:

La Ingeniería de Software dispone de un cuerpo de conocimiento que permite definir modelos de proceso para desarrollar artefactos software. Sin embargo, la naturaleza de los EVT's, caracterizados como facilitadores de interacción entre humanos físicamente no contiguos, y la orientación a personas con discapacidades diferentes requiere la definición de procesos específicos para la construcción, el diseño y la evaluación que tengan presente ese aspecto.

Objetivo General:

Los espacios virtuales de trabajo constituyen una familia emergente de aplicaciones a desarrollar, por lo que necesita un enfoque ingenieril. También existe la necesidad de incluir en estos desarrollos a personas con distintas discapacidades, de tal forma que un no vidente o un hipoacúsico puedan interactuar por medio de estos espacios virtuales sin que su discapacidad signifique una barrera. El presente proyecto busca desarrollar un proceso de diseño de dispositivos educativos y evaluación basado en Espacios Virtuales de Trabajo y que atiendan de manera integral las necesidades de personas con distintas discapacidades cumpliendo con la normativa vigente.

Objetivos Específicos:*Objetivo Específico 1:*

Desarrollar un Proceso de Diseño de Dispositivos Educativos Centrados en Espacios Virtuales de Trabajo y orientado a personas con discapacidad física.

Objetivo Específico 2:

Desarrollar un Conjunto de Dispositivos Educativos Centrados en Espacios Virtuales de Trabajo y orientado a personas con discapacidad física, utilizando el proceso anterior, para implementar en el ámbito de la UNLa.

Objetivo Específico 3:

Desarrollar un Proceso de Evaluación de Espacios Virtuales de Trabajo Orientados a Educación y orientado a personas con discapacidad física, e implementarlo en el ámbito de la UNLa.

Objetivo Específico 4:

Implementar el Proceso de Evaluación de Espacios Virtuales de Trabajo Orientados a Educación y orientado a personas con discapacidad física, en el ámbito de la UNLa.

METODOLOGÍA DE TRABAJO

Para construir el conocimiento asociado al presente proyecto de investigación, se seguirá un enfoque de investigación clásico [Riveros y Rosas, 1985; Creswell, 2002] con énfasis en la producción de tecnologías [Sábato y Mackenzie, 1982]; identificando métodos, materiales y abordaje metodológico necesarios para desarrollar el proyecto.

Métodos:*Revisiones Sistemáticas:*

Las revisiones sistemáticas [Argimón, 2004] de artículos científicos siguen un método explícito para resumir la información sobre determinado tema o problema. Se diferencia de las revisiones narrativas en que provienen de una pregunta estructurada y de un protocolo previamente realizado.

Prototipado Evolutivo Experimental (Método de la Ingeniería):

El prototipado evolutivo experimental [Basili, 1993] consiste en desarrollar una solución inicial para un determinado problema, generando su refinamiento de manera evolutiva por prueba de aplicación de dicha solución a casos de estudio (problemáticas) de complejidad creciente. El proceso de refinamiento concluye al estabilizarse el prototipo en evolución.

Materiales:

Para el desarrollo de los formalismos y procesos propuestos se utilizarán:

- Formalismos de modelado conceptual usuales en la Ingeniería de Software [Rumbaugh et al., 1999; Jacobson et al., 2013] y en la Ingeniería del Conocimiento [García-Martínez y Britos, 2004].

- Modelos de Proceso usuales en Ingeniería de Software [IEEE, 1997; ANSI/IEEE, 2007; Oktaba et al., 2007].
- Normativas vigentes sobre las responsabilidades del Estado [Ley Educación Superior, 2002; Ley Accesibilidad Web, 2010].
- Normas de Accesibilidad Web [W3C, WAI].

RESULTADOS OBTENIDOS/ESPERADOS

Se espera tener las versiones de los siguientes productos: (i) Proceso de Diseño de Dispositivos Educativos Centrados en EVT's y orientado a personas con discapacidad, (ii) Conjunto de Dispositivos Educativos integrable a etapas de producción de EVT's Educativos y orientado a personas con discapacidad implementados en la UNLa, (iii) Proceso de Evaluación de EVT's Orientados a Educación y orientado a personas con discapacidad, y (iv) la Implementación del Proceso de Evaluación de EVT's integrable a etapas de producción de EVT's Educativos y orientado a personas con discapacidad en el ámbito de la Universidad Nacional de Lanús.

Se prevé publicar resultados en las siguientes Series y Revistas:

- Lecture Notes on Artificial Intelligence (Springer Verlag)
- Lecture Notes on Computer Science (Springer Verlag)
- Revista Latinoamericana de Ingeniería de Software

Y en los siguientes Eventos Científicos:

- Workshop de Investigadores en Ciencias de la Computación
- Latin American Congress on Requirements Engineering & Testing
- Latin American Symposium on Software Engineering
- Congreso Argentino de Ciencias de la Computación
- Congreso Argentino de Tecnología en Educación y Educación en Tecnología
- Seminario Internacional de Educación a Distancia de RUEDA
- Frontiers in Education
- Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento

FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo está conformado por dos investigadores formados y cuatro en formación. En

su marco se desarrollan dos trabajos finales de licenciatura.

FINANCIAMIENTO

Las investigaciones que se proponen en esta comunicación cuentan con financiamiento como Proyecto 80020170300059LA de la Secretaría de Ciencia y Técnica de la Universidad Nacional de Lanús (Argentina).

REFERENCIAS

- ANSI/IEEE, (2007). *Draft IEEE Standard for software and system test documentation*. ANSI/IEEE Std P829-2007.
- Argimón J. (2004). *Métodos de Investigación Clínica y Epidemiológica*. Elsevier España, S.A. ISBN 9788481747096.
- Basili, V. (1993). *The Experimental Paradigm in Software Engineering*. En *Experimental Software Engineering Issues: Critical Assessment and Future Directions* (Ed. Rombach, H., Basili, V., Selby, R.). Lecture Notes in Computer Science, Vol. 706. ISBN 978-3-540-57092-9.
- Charum, V. (2007). *Modelo de Evaluación de Plataformas Tecnológicas*. Tesis de Magíster en Telecomunicaciones. Escuela de Posgrado, Instituto Tecnológico de Buenos Aires.
- Creswell, J. (2002). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Prentice Hall. ISBN 10: 01-3613-550-1.
- Ferreira Szpiniak, A., Sanz, C. (2007). *Hacia un modelo de evaluación de entornos virtuales de enseñanza y aprendizaje*. *Proceedings XIII Congreso Argentino de Ciencias de la Computación*. Pag. 932-947. ISBN 950-656-109-3.
- Ferreira Szpiniak, A., Sanz, C. (2012). *MUSA un modelo de evaluación de Entornos virtuales de enseñanza y aprendizaje. Aplicación a un caso de estudio*. Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología. TE&ET. ISSN, 9959, 94-103.
- García Martínez, R., Britos, P. (2004). *Ingeniería de Sistemas Expertos*. Editorial Nueva Librería. ISBN 987-1104-15-4.
- Gibson, C., Cohen, S. (Eds.). (2003). *Virtual Teams That Work: Creating Conditions for Virtual Team Effectiveness*. John Wiley & Sons.
- González, A. H., Esnaola, F., Martín, M. (2012). *Propuestas educativas mediadas por tecnologías digitales*. Dirección de Educación a Distancia, Innovación en el Aula y TIC (EAD). Universidad Nacional de La Plata.
- Grangel, R., Campos, C., Rebollo, C., Remolar, I., Palomero, S. (2012). *Metodología para seleccionar tecnologías Web 2.0 para la docencia*. *Jornadas de Enseñanza de la Informática (18es: 2012: Ciudad Real)*.
- Herrera, A., Rodríguez, D., García Martínez, R. (2014). *Topologías de Interacción para Espacios Virtuales de Trabajo Colaborativo*. Revista Iberoamericana de Educación en Tecnología y Tecnología en Educación, 14: 74-86. ISSN 1850-9959.
- Herrera, A., Rodríguez, D., García-Martínez, R. (2013). *Taxonomía de Mecanismos de Awareness*. Actas del XI Workshop de Tecnología Informática Aplicada en Educación. *Proceedings XIX Congreso Argentino de*

- Ciencias de la Computación. Pag. 651-660. ISBN 978-987-23963-1-2.
- IEEE, (1997). *IEEE Standard for Developing Software Life Cycle Processes. IEEE Std 1074-1997* (Revisión of IEEE Std 1074-1995; Replaces IEEE Std 1074.1-1995)
- Jacobson, I., Ng, P. W., McMahon, P. E., & Jaramillo, C. M. Z. (2013). *La esencia de la ingeniería de software: El núcleo de Semat*. Revista Latinoamericana de Ingeniería de Software, 1(3), 71-78.
- Lazar, J.; Dudley, A.; Sponaugle, K. (2004). *Computers in Human Behavior Improving Web accessibility: a study of webmaster perceptions*. Editorial Elsevier.
- Ley de Educación Superior (2002). Ley Nacional Nro. 25573 (modifica Ley 24521). Publicada en el Boletín Oficial del 30-abr-2002, Número: 29888, Página: 1
- Ley de Accesibilidad Web (2010). Ley 26.653 Accesibilidad de la Información en las Páginas Web. Autoridad de Aplicación. Plazos. Reglamentación. Sancionada: Noviembre 3 de 2010. Promulgada de Hecho: Noviembre 26 de 2010
- Marqués Graells, P. (1999). *Criterios para la clasificación y evaluación de espacios web de interés educativo*. Educar, (25): 095-111.
- Mikropoulos, T., Natsis, A. (2011). *Educational Virtual Environments: A Ten-Year Review of Empirical Research (1999–2009)*. Computers & Education, 56(3); 769-780.
- Monereo, C., Coll, C., 2008, *Psicología de la educación virtual*, Ediciones Morata, Madrid.
- Oktaba, H., Garcia, F., Piattini, M., Ruiz, F., Pino, F., Alquicira, C. (2007). *Software Process Improvement: The Competisoft Project*. IEEE Computer, 40(10): 21-28. ISSN 0018-9162.
- Paciello, M.(2000). *Web Accessibility for People with Disabilities*. Series R&D Developers.
- Redfern, S., & Naughton, N. (2002). *Collaborative Virtual Environments to Support Communication and Community in Internet-Based Distance Education*. Journal of Information Technology Education, 1(3).
- Riveros, H. y Rosas, L. (1985). *El Método Científico Aplicado a las Ciencias Experimentales*. Editorial Trillas. México. ISBN 96-8243-893-4.
- Rumbaugh, J., Jacobson, I., Booch, G. (1999). *The Unified Modeling Language, Reference Manual*. Addison Wesley, ISBN-10: 02-0130-998-X.
- Sabato J, Mackenzie M. (1982). *La Producción de Tecnología: Autónoma o Transnacional*. Instituto Latinoamericano de Estudios Transnacionales - Technology & Engineering. ISBN 9789684293489.
- Schneiderman, Ben (2006). *Access by Design*. New Riders Press
- Tatcher, J.; Burks, M.; Heilmann, C. (2006). *Web Accessibility: Web Standards and Regulatory Compliance*.
- W3C: Sitio oficial de la W3C. Disponible en <http://w3.org>
- WAI: Sitio oficial de la Iniciativa WAI de la W3C. Disponible en <http://w3.org/WAI>

Pensamiento Computacional y Programación en la Formación de Docentes del Nivel Primario

Ana Casali^{1,2}, Dante Zanarini^{1,2}, Patricia San Martín³, Natalia Monjelat³

¹ Facultad de Ciencias Exactas, Ingeniería y Agrimensura,
Universidad Nacional de Rosario, Rosario, Argentina.
{acasali, dante}@fceia.unr.edu.ar

² Centro Internacional Franco Argentino de Ciencias de la Información y de Sistemas
(CIFASIS: CONICET-UNR)

³ Instituto Rosario de Investigaciones en Ciencias de la Educación
(IRICE: CONICET-UNR)
{sanmartin, monjelat}@irice-conicet.gov.ar

RESUMEN

En este trabajo se presenta una línea de investigación orientada al desarrollo de dispositivos para la formación docente del nivel primario en Pensamiento Computacional y Programación, combinando el enfoque desconectado con actividades que impliquen el uso de computadoras y otros dispositivos. Como una primera instancia de formación se presenta el diseño y la implementación de la “Especialización Docente de Nivel Superior en Didáctica de las Ciencias de la Computación”. El objetivo de esta especialización es que los docentes de primaria puedan apropiarse de esta forma de pensamiento y del potencial de la programación, para incluirlas en sus prácticas de manera contextualizada, interdisciplinar e inclusiva.

Palabras clave: Pensamiento Computacional, Programación, Formación docente, Educación primaria.

CONTEXTO

Esta línea de I+D se está llevando a cabo en un grupo interdisciplinario a través de varios proyectos que comparten alguno de sus

miembros docentes e investigadores, conformando una red de colaboración:

- Especialización Docente de Nivel Superior en Didáctica de las Ciencias de la Computación: Aprendizaje y Enseñanza del Pensamiento Computacional y la Programación en el Nivel Primario, A. Casali, D. Zanarini, P. San Martín, N. Monjelat, C. Ortega y A. Baldomá. Convenio Fundación Sadosky, UNR y Ministerio de Educación Pcia. Santa Fe (2016-2019).

- Proyecto de Investigación Científica y Tecnológico (PICT) N° 1530: “Aprender a programar en primaria: Hacia la construcción de Tecnologías para la inclusión social”, N. Monjelat (2017-2019).

- Desarrollo del Pensamiento Computacional a través de actividades desconectadas en la educación básica D. Barone, A. Casali y C. Puhlmann Brackmann UFRGS, Brasil (2016-2017).

1. INTRODUCCIÓN

La introducción de nociones relacionadas con las Ciencias de la Computación (CC) en el currículum escolar, especialmente en cuanto

al desarrollo del pensamiento computacional (PC) y la apropiación de conceptos de programación, se ha ponderado en las últimas décadas con un importante consenso en el campo educativo internacional. Siguiendo a Wing [7] el aprendizaje de CC beneficia a toda la sociedad y por lo tanto, su enseñanza debería estar presente en todos los niveles educativos.

Argentina lanzó en el año 2013 el proyecto “Program.ar”, que a partir de múltiples iniciativas acerca a niños y jóvenes al aprendizaje de las CC. Asimismo, el Consejo Federal de Educación señaló al aprendizaje de la programación de importancia estratégica para el Sistema Educativo Nacional durante la escolaridad obligatoria (Resolución CFE N° 263/15, 2015) afirmando la relevancia de estos contenidos para la formación de los niños y jóvenes. Sin embargo en el marco institucional, aún no se han efectivizado dos aspectos clave interrelacionados que condicionan fuertemente la introducción de nociones de las CC en el sistema educativo argentino. Uno de estos aspectos refiere a la necesidad de realizar modificaciones en los planes de estudio tanto a nivel primario y secundario, así como en la formación docente. El otro, implica la inmediata puesta en obra de propuestas de formación docente en el aprendizaje y enseñanza del pensamiento computacional y la programación que habiliten paulatinas transformaciones tanto de lo curricular como de la práctica educativa escolar [5]. En particular, en el contexto de la provincia de Santa Fe, se observa que casi todas las escuelas públicas e institutos de formación docente fueron dotados de infraestructura tecnológica digital a través de programas nacionales y provinciales. Por ejemplo, en el nivel primario se cuenta con las “Aulas digitales móviles” provenientes tanto del programa nacional “Primaria Digital” como del provincial “Tramas Digitales”. En referencia a la formación y capacitación docente, la misma generalmente se ha enfocado hacia un primer nivel de apropiación

instrumental de aplicaciones, herramientas y recursos digitales de las actuales Tecnologías de la Información y Comunicación (TIC) vinculadas a la práctica educativa. En este sentido, un relevamiento realizado en agosto de 2016, concluyó que ninguno de los 139 Institutos de Formación Docente de la provincia ofrecía un postítulo sobre la didáctica de las CC para la formación de profesores. Por ello, es posible afirmar que la presente vacancia impacta en la formación de todos los estudiantes y en el logro de una articulación satisfactoria entre los distintos niveles del sistema educativo, así como sobre el número de egresados de nivel terciario y universitario en áreas estratégicas para el desarrollo regional.

A partir de esta necesidad y considerando los lineamientos del Ministerio de Educación de Santa Fe que proponen el desarrollo de proyectos institucionales en relación a los Núcleos Interdisciplinarios de Contenidos, se considera posible llevar adelante una especialización docente de nivel superior en didáctica de las CC, desde una doble articulación que considere tanto los procesos como los productos implicados en el desarrollo del pensamiento computacional y la programación, como posibles Tecnologías para la Inclusión Social [4]. Por otra parte, los trabajos realizados en distintos contextos de educación primaria, han demostrado que el desarrollo de procesos de enseñanza y aprendizaje a través de actividades lúdicas “unplugged” o desconectadas, sin uso de computadora, resulta efectivo para desarrollar habilidades cognitivas de pensamiento computacional [1] [2] [3].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Dada la necesidad de formación docente en el área, se planteó el diseño de la Especialización Docente de Nivel Superior en Didáctica de las

Ciencias de la Computación: Aprendizaje y Enseñanza del Pensamiento Computacional y la Programación en el Nivel Primario. Esta formación tiene por objetivo general formar docentes capaces de experimentar y reflexionar críticamente acerca de los procesos de desarrollo del PC y la programación, a los fines de construir las competencias adecuadas al nivel primario que posibiliten una práctica educativa innovadora con énfasis en la resolución de problemas mediante la producción colaborativa e interdisciplinaria de Tecnologías para la Inclusión Social.

En referencia a los objetivos específicos, la presente propuesta de formación especializada pretende:

- Promover el estudio de los principios del pensamiento computacional y la programación para su adecuación didáctica al nivel primario de escolaridad.
- Activar un posicionamiento crítico y ético acerca del uso, impacto y potencial de las TIC en el contexto socio-cultural actual.
- Brindar herramientas teóricas, metodológicas y técnicas para la selección y aplicación de conceptos del PC y la programación en función del diseño y desarrollo de proyectos educativos enfocados hacia la producción de tecnologías inclusivas.
- Desarrollar el interés y compromiso responsable hacia la participación en la producción colaborativa de programas sencillos integrados tanto a temáticas de la educación primaria como a proyectos institucionales utilizando diversas herramientas, bajo metodologías de trabajo interdisciplinario.
- Favorecer en el marco del sistema educativo provincial, el desarrollo de procesos institucionales de adecuación curricular y transformación de las prácticas educativas de nivel primario aportando fundamentos teórico-metodológicos y técnicos sobre contenidos relacionados a las CC.

Diseño Curricular

La Especialización se organizó en cuatro semestres, presentando en cada uno de ellos tres módulos (doce módulos en total), en los que se articulan los contenidos de manera espiralada. Los módulos se desarrollaron sobre los ejes del pensamiento computacional, la programación y conceptos tecnológicos. La propuesta incluye también proyectos integradores, con el fin de que faciliten a los docentes cursantes llevar al aula estos procesos de enseñanza y aprendizaje planificando y desarrollando actividades no excluyentes adecuadas a su contexto institucional. Cada semestre tiene asignada una carga horaria de 100 hs. de dictado totalizando 400 hs., donde el 80% corresponde a actividades presenciales bajo el formato de taller y seminario-taller, como se muestra en la Tabla I.

Tabla I: Especialización: Distribución de Módulos y carga horaria

Sem.	Módulo	Modalidad	Hs
1	Introducción: CC en la Escuela Primaria	Seminario-Taller	20
	Pensamiento Computacional I	Seminario-Taller	40
	Taller de Programación I	Taller	40
2	Proyecto Integrador I	Taller	20
	Pensamiento Computacional II	Seminario-Taller	40
	Taller de Programación II	Taller	40
3	Proyecto Integrador II	Taller	20
	Introducción a la Organización de Computadoras	Seminario-Taller	40
	Taller de Robótica	Taller	40
4	Ciencia de datos, privacidad y seguridad de la información	Seminario-Taller	35

3

	Desarrollo de aplicaciones móviles	Taller	40
	Proyecto Final	Taller	25
Total	12 Módulos		400

Las actividades presenciales se continúan y complementan con actividades virtuales realizadas a través de la Plataforma Educativa de la provincia de Santa Fe (<http://plataformaeducativa.santafe.gov.ar/>), donde los docentes son usuarios registrados, accediendo tanto a prestaciones administrativas como a otros cursos de capacitación del Ministerio de Educación.

3. RESULTADOS

El postítulo fue aprobado por el Ministerio de Educación provincial en julio de 2017 y comenzó a dictarse en agosto del mismo año, en el ISFD N°36 "Mariano Moreno" de la ciudad de Rosario. El equipo interdisciplinario a cargo del dictado estuvo compuesto por cuatro profesores del Departamento de Ciencias de la Computación (UNR), dos profesoras del ISFD y dos docentes-investigadores del IRICE. A partir de la publicidad del mismo se obtuvieron 120 inscriptos de los cuales 89 se presentaron al dictado. De este grupo, el 70% cumplió con la asistencia requerida. Actualmente se encuentran en ejecución una serie de instrumentos para obtener información respecto al perfil más completo de los docentes cursantes, recopilar sus opiniones sobre el postítulo en general y el dictado de cada módulo.

Desde el punto de vista de los profesores que participaron del dictado, las actividades se mantuvieron dentro de lo planificado. Se realizaron las adecuaciones necesarias a la dinámica del aprendizaje del grupo, considerando siempre el trabajo en modalidad de taller con actividades lúdicas que ellos puedan contextualizar en sus aulas. En general, el abordaje seguido en los distintos encuentros

fue: 1) breve presentación de conceptos básicos y teóricos sobre el tema a trabajar, luego 2) planteo de una serie de actividades (ejercicios con enfoque lúdico adecuados para el nivel primario) de dificultad creciente a desarrollar en forma grupal y como cierre 3) puesta en común de los resultados alcanzados grupalmente, orientando las reflexiones para llegar a los objetivos que fueron planteados en la actividad. En el transcurso de los encuentros, ha resultado más efectivo intensificar las actividades prácticas con integraciones conceptuales y profundizar el cierre reflexivo.

Dada la diversidad de especialidades de los docentes, los trabajos prácticos desarrollados se han visto enriquecidos por la diversidad y la integración de saberes. La actividad de cierre de la especialización consistió en un taller donde los distintos grupos presentaron sus trabajos finales de PC y de programación. Luego hubo un espacio de reflexión grupal donde los comentarios fueron muy satisfactorios respecto a los avances logrados y la percepción de los cursantes considerando los logros alcanzados.

Actualmente se encuentra en evaluación el primer semestre de la propuesta de formación planteada desde el punto de vista de los aprendizajes alcanzados por los cursantes y de sus opiniones sobre el dictado. Como observaciones preliminares todos los docentes han valorado el trabajo en pensamiento computacional de forma desconectada (unplugged) ya que posibilita incluir actividades que promuevan el aprendizaje de estos conceptos y habilidades, en todos los contextos, contando con los recursos que están disponibles en todas las escuelas. Además, en cada módulo se han planteado actividades que pueden ser incluidas transversalmente y en vinculación con otros contenidos de la educación primaria. Por otra parte, los avances en programación han sido muy gratificantes, sorprendiendo principalmente a los

participantes por los logros alcanzados.

A partir de los resultados, aprendizajes y evaluación de esta experiencia se espera realizar un intercambio con otras instancias de formación similares en distintas regiones, a fin de difundir la experiencia y crear redes que fortalezcan a las CC en el contexto de educación primaria tanto en el país como en latinoamérica.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo es interdisciplinario y está integrado por las doctoras Ana Casali, Patricia San Martín, Natalia Monjelat y el licenciado Dante Zanarini, con el apoyo de la Rectora del ISFD Claudia Ortega y la Profesora de dicha institución Alejandra Ortega. Dentro del marco de esta línea de I+D se desarrollan actualmente las siguientes tesis doctorales:

- Christian Brackmann: “Desenvolvimento do Pensamento Computacional Através de Atividades Desplugadas na Educação Básica”, Dir. D. Barone, Co-dirección A. Casali, UFRGS, Brasil 2017.

- Marisa Cenacchi: “Hacia una Educación no excluyente. Perspectivas teóricas-metodológicas sobre “Accesibilidad-DHD” para la formación superior de educadores”. Dir. P. San Martín, Co-dir. G. Guarnieri, UNR, Argentina (en evaluación).

Cabe destacar la formación que se lleva a cabo de docentes del Depto. de Ciencias de la Computación en la enseñanza de la disciplina a docentes del nivel primario, en un trabajo interdisciplinario con especialistas de las ciencias de la educación.

5. REFERENCIAS

[1] Bell, Witten & Fellows. (2015). CS Unplugged: An enrichment and extension programme for primary-aged students.

[2] Brackmann, C., Barone, D., Casali, A., Boucinha, R., Muñoz-Hernandez, S. Computational Thinking: Panorama of the Americas, SIIE 2016. Salamanca, España. Disponible

<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7751839>

[3] C. Brackmann, M. Román-González, J. Moreno-León, G. Robles, A. Casali and D. Barone, Computational Thinking Unplugged: Teaching and Student Evaluation in Primary Schools. In Proceedings WIPSCE Nijmegen, The Netherlands, ACM. November, 2017. <https://dl.acm.org/citation.cfm?id=3137069&CFID=830581293&CFTOKEN=20769223>

[4] Monjelat, N., & San Martín, P. (2016). Programar con Scratch en contextos educativos: ¿Asimilar directrices o co-construir Tecnologías para la Inclusión Social? Praxis Educativa, 20(1), 61–71. Disponible en: <http://ojs.fchst.unlpam.edu.ar/ojs/index.php/praxis/article/viewFile/1051/1092>

[5] San Martín, P. y Guisen, M. A. (2016) Hacia las tecnologías para la inclusión social en contextos educativos regionales: análisis del caso “ECCA”. Actualidades Investigativas en Educación, 16(2), 1-26.

[6] Thomas, H., Juarez, P. y Picabea, F. (2015) ¿Qué son las tecnologías para la inclusión social? 1º Edición. Bernal: Universidad de Quilmes. Disponible en: https://issuu.com/redtisa/docs/cuadernillo_n1_online

[7] Wing, Jeannette (2014) Computational Thinking Benefits Society. Social Issues in Computing. <http://socialissues.cs.toronto.edu/2014/01/computational-thinking/>

[8] N. Monjelat, (2017) “Programming Technologies for Social Inclusion,” En LACLO, 2017. La Plata, Argentina. Disponible en: <http://ieeexplore.ieee.org/document/8120901/>

Entorno basado en programación por bloques para robots educativos

Juan de la Fuente Jorge Rodríguez Rafael Zurita Laura Cecchi

email: {juan.delafuente, j.rodrig, rafa, lcecchi}@fi.uncoma.edu.ar

Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

Resumen

Existen diversos programas, iniciativas y proyectos orientados a promover la enseñanza y el aprendizaje de la programación en el ámbito de la educación obligatoria.

En este marco se plantea como objetivo el desarrollo de iniciativas orientadas a la construcción de conocimiento y recursos didácticos para la enseñanza de la programación en el ámbito de la robótica educativa.

Este trabajo presenta una Línea de Investigación y Desarrollo que extiende al Proyecto Scratch sumando funcionalidades para interactuar con el robot educativo Frankestito, desarrollado en la Facultad de Informática de la Universidad Nacional del Comahue, a fin de obtener un entorno basado en programación por bloques para robots educativos. Se espera que este entorno mejore las oportunidades de aprender conceptos fundamentales de las Ciencias de la Computación a una población más amplia de estudiantes, penetrando sólidamente en la educación obligatoria.

Palabras Clave: ENSEÑANZA DE LA PROGRAMACIÓN, ENSEÑANZA DE LAS CIENCIAS DE LA COMPUTACIÓN, RECURSOS DIDÁCTICOS EN CIENCIAS DE LA COMPUTACIÓN, EDUCACIÓN OBLIGATORIA, PROGRAMACIÓN BASADA EN BLOQUES, ROBÓTICA EDUCATIVA.

Contexto

Esta propuesta se ubica en el contexto de las iniciativas promovidas por el Grupo de Investigación en Lenguajes e Inteligencia Artificial de la Facultad de Informática y del Convenio Marco de Colaboración firmado durante 2016 entre la Facultad de Informática y el Ministerio de Educación de la Provincia del Neuquén.

Este trabajo se desarrolla en el ámbito de los proyectos de investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación (04/F015)* y *Agentes Inteligentes y Web Semántica (04/F014)* financiados por la Universidad Nacional del Comahue a través de la Secretaría de Ciencia y Técnica y avalados por el Consejo Provincial de Educación en el contexto del Convenio Marco de Colaboración. Los proyectos tienen prevista una duración de cuatro años a partir de enero del 2017.

1. Introducción

Actualmente, las nuevas tecnologías atraviesan transversalmente toda la sociedad en menor o mayor grado. Sin embargo, con frecuencia nos adentramos en actividades virtuales únicamente desde un rol de “consumidor” de ellas, desconociendo su concepto básico intrínseco. Bajo ésta óptica urge incor-

porar conceptos de las Ciencias de la Computación a la educación formal brindando herramientas que permitan no sólo “leer” tecnología, sino ser capaces de “escribirla”[4, 11, 12].

En este sentido, en diversos países de Europa y en Estados Unidos, organizaciones gubernamentales[15, 3] y no gubernamentales[8, 16] han advertido sobre la necesidad de reformar drásticamente la enseñanza actual en las áreas de *Informática* y *Computación* en la educación obligatoria para introducir conceptos vinculados a las Ciencias de la Computación.

Haciendo eco de estas iniciativas a nivel nacional, el Consejo Federal de Educación declaró de importancia estratégica la enseñanza y el aprendizaje de la Programación en la escolaridad obligatoria [5]. A partir de esta declaración y de los proyectos nacionales ya existentes, tales como la iniciativa Program.ar y el Programa Nacional PLANIED, la necesidad de abordar la enseñanza de estos temas en el ámbito de la escuela es clara y prioritaria.

En este trabajo se presenta una Línea de Investigación que busca desarrollar vínculos para potenciar dos herramientas ya existentes que se han posicionado como efectivas para favorecer la construcción de este tipo de conocimiento en la educación primaria y secundaria: El proyecto *Scratch* y *Frankestito*.

Proyecto Scratch

Investigadores del MIT (Massachusetts Institute of Technology) idearon un Framework multiplataforma de desarrollo de aplicaciones y pensamiento computacional, diseñado para ser altamente interactivo con una gramática basada en una colección de “bloques de programación” (gráficos apilables) que permiten crear programas sin tipear código, llamado *Scratch* [12]. Este proyecto se encuentra vigente y utilizado en más de 150 países, Argentina entre ellos[7].

Frankestito

En el ámbito regional, la Facultad de In-

formática (FAIF-UNCo) desarrolló un robot educativo, denominado Frankestito[9], para la enseñanza y promoción de contenidos relacionados con las Ciencias de la Computación en distintos niveles educativos[13]. El mismo cuenta con una serie de características distintivas, tales como, ser de bajo costo, contar con una conexión inalámbrica y persistente, ser myro-compatible y poseer visión del ambiente. Esto lo convierte en un sujeto propicio para el desarrollo de herramientas que potencien y diversifiquen su funcionalidad.

Ambas herramientas han logrado tener una buena recepción en sus ámbitos de influencia [10, 13, 14], sin embargo existen potencialidades aún no desarrolladas en ellas.

Por una parte *Frankestito* soporta una diversidad de lenguajes clásicos (Python, Prolog, C, C++, entre otros) que permiten interactuar de forma efectiva con el robot educativo, pero no cuenta con un entorno propicio que facilite su utilización sin intervención de líneas de código, limitando su llegada a otros rangos etarios o ámbitos de aplicación educativa.

Mientras que *Scratch* provee un Framework para programación basada en bloques fácilmente utilizable sin necesidad de *escribir líneas de código*.

En esta línea de trabajo proponemos extender *Scratch* para comunicarse con robots compatibles con la arquitectura de Frankestito, potenciándolos al permitir incorporar un lenguaje y entorno de desarrollo basado en bloques para su utilización.

La idea es añadir a *Scratch*, como nuevos bloques de programación, todas las funciones con las que cuenta Frankestito a fin de ser utilizados por los estudiantes o *scratchers*, trabajando en el entorno del Framework, para controlar el robot, complementando el lenguaje de scripting tradicional. De este modo, se logrará un entorno de programación sencillo y transparente con comunicación efectiva y simplificada.

Cabe aclarar que aunque Frankestito actualmente puede ser programado en diferen-

tes lenguajes de programación, no cuenta con un entorno de programación basada en bloques. Existen desarrollos que extienden *Scratch* enfocados en la robótica, tales como LEGO WeDo[1] y PicoBoard[2]. Sin embargo, los mismos no se adaptan a las necesidades de Frankestito, dado que cuentan con una funcionalidad notablemente más limitada.

La estructura del presente trabajo es la siguiente. En la sección 2 presentamos los objetivos de los proyectos de investigación en los que se enmarca este trabajo y describimos la línea de investigación actual. En la sección 3 indicamos algunos resultados obtenidos y trabajos futuros. Finalmente, comentamos aspectos referentes a la formación de recursos humanos en esta temática.

2. Línea de investigación y desarrollo

El proyecto de investigación *Agentes Inteligentes y Web Semántica* tiene entre sus objetivos generar conocimiento especializado en el área de agentes inteligentes. Particularmente se trabaja en la programación de robots para generar comportamiento inteligente. Por otra parte el proyecto de investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación*, tiene como uno de sus objetivos el desarrollo de modelos que contribuyan a la producción de un marco teórico que se ocupe de estudiar la inclusión de la computación en la educación y el diseño e implementación herramientas que soporten estos modelos.

Ambos proyectos confluyen en esta línea de investigación que propone la integración de un entorno de programación basado en bloques a la plataforma de robots educativos compatibles con Frankestito. Los resultados de este trabajo, permitirán ampliar las posibilidades didácticas de Frankestito e involucrar en la enseñanza y aprendizaje de la programación a una nueva población de docentes y estudiantes.

Por esto, se desarrollará una extensión que permita conectar el *Proyecto Scratch* con la

plataforma de robots educativos *Frankestito*.

Bajo esta premisa se establecen los siguientes ejes de trabajo:

Scratch

Recientemente el Proyecto Scratch incorpora *Scratch Extensions* que permite sumar colecciones de bloques destinados interactuar con dispositivos externos [6].

Esta incorporación, que resalta la necesidad de contar con nuevos recursos didácticos, facilita la posibilidad de potenciar al Framework con complementos tangibles que interactuen con los productos de software logrados por los *Scratchers*.

Para ello en este eje de trabajo se desarrollará una colección de bloques fácilmente manipulables, que cumplan con las especificaciones del *Proyecto Scratch* y concentren, en una abstracción de código, las funcionalidades para interactuar con robots educativos de la arquitectura *Frankestito*. Esto es factible gracias a su licencia *GPLv2* y *Scratch Source Code License*.

Frankestito

Frankestito resulta un candidato ideal para investigación al haber demostrado ser compatible con otras plataformas de robótica educativa, como el Multiplo N6-Max[17]. De esta manera, cualquier desarrollo que sea factible sobre la arquitectura Frankestito es transferible a otras plataformas.

Este eje de trabajo tiene por objeto desarrollar herramientas que permitan ampliar los horizontes de su utilización, brindando un entorno que lo posicione como accesible para cualquier nivel del sistema educativo.

Para ello se desarrollará un protocolo que facilite el envío de órdenes sobre su arquitectura de comunicación inalámbrica, otorgando transparencia en el intercambio de mensajes y desvinculada de algunas configuraciones previas requeridas para su utilización, para lograr una independencia del lenguaje de scripting nativo del robot.

Plataforma de Comunicación

Se considera necesario diseñar e implementar

una capa de comunicación entre las extensiones desarrolladas para *Scratch* y el protocolo para envío de órdenes y recepción de respuestas construido sobre la plataforma *Frankestito*.

Este eje de trabajo se enfoca en el diseño de una capa de comunicación vía HTTP/Socket bajo la implementación de un protocolo que permita estandarizar estas comunicaciones de modo que puedan ser fácilmente adaptables para otros entornos de programación o frameworks y compatibles con la plataforma de robots de esta arquitectura.

Estos tres ejes de trabajo no se presentan de forma aislada; se considera que la investigación en estas áreas debe plantearse como integrada y estudiarse en forma conjunta para avanzar en el desarrollo de un producto de software consistente que permita la continuidad en la investigación de nuevas herramientas. Todo resultado conservará patentes compatibles con los proyectos relacionados.

3. Resultados obtenidos y Trabajos Futuros

Inicialmente, se analizó la arquitectura de la plataforma *Frankestito* y la adaptación del robot Multiplo N6 Max[17], robot compatible con *Frankestito*, identificando funcionalidades específicas a ser implementadas para lograr la interacción desde el entorno de *Scratch* que extenderán el lenguaje de programación. Se hizo énfasis en aquellas que conciernen a la comunicación.

Actualmente, se está realizando una revisión sobre diferentes entornos de programación por bloque, a fin de identificar las características principales de este tipo de lenguajes de programación. El estudio está centrado en *Scratch*.

A partir de estos resultados se realizará el diseño de una arquitectura que adapte y extienda el entorno de programación *Scratch*, con bloques correspondientes a la interacción

con *Frankestito*. En esta actividad se considerarán los comentarios, sugerencias y revisiones reunidas de experiencias de enseñanza y aprendizaje en entornos reales, realizadas sobre el trabajo de campo. El desarrollo de esta investigación será documentado en un reporte técnico.

En el contexto de esta Línea de Investigación y Desarrollo se espera obtener los siguientes resultados:

- Dotar a *Scratch* de todas las funciones con las que cuenta *Frankestito*, como nuevos bloques de programación, a fin de ser utilizados por estudiantes, docentes o *scratchers*.
- Desarrollar una capa de comunicación entre la plataforma *Frankestito* que permita la interacción con *Scratch* y otros entornos de programación.
- Introducir mejoras en la plataforma *Frankestito* tendientes a favorecer la interacción con la capa de comunicación,

4. Formación de Recursos Humanos

Sobre la temática de esta línea de investigación, uno de los autores de este trabajo está desarrollando su tesis de grado de la Licenciatura en Ciencias de la Computación.

Por otra parte, otro de los autores de este trabajo está inscripto en la Maestría en Enseñanza en Escenarios Digitales que desarrollan de manera conjunta las Universidades Nacionales de Cuyo, Comahue, Patagonia Austral, Patagonia San Juan Bosco, San Luis, Chilecito y La Pampa.

Referencias

- [1] Sitio Oficial de *Scratch* Wiki: Lego WeDo. Último acceso Marzo 2018, website https://wiki.scratch.mit.edu/wiki/LEGO_Education_WeDo_Robotics_Kit.

- [2] Sitio web oficial de Scratch Wiki PicoBoard. Último acceso Marzo 2018, website <https://wiki.scratch.mit.edu/wiki/PicoBoard>.
- [3] S. Bocconi, A. Chiocciariello, G. Dettori, A. Ferrari, K. Engelhardt, P. Kampylis, and Y. Punie. Developing computational thinking in compulsory education. *European Commission, JRC Science for Policy Report*, 2016.
- [4] M. Borchardt and I. Roggi. Ciencias de la computación en los sistemas educativos de américa latina. 2017.
- [5] Consejo Federal de Educación. Res 263/15. *Resoluciones CFE*, 2015.
- [6] S. Dasgupta, S. M. Clements, A. Y. Idlbi, C. Willis-Ford, and M. Resnick. Extending scratch: New pathways into programming. In *2015 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pages 165–169, Oct 2015.
- [7] P. Factorovich and F. Sawady O'Connor. *Actividades para aprender a Program.AR*. Fundación Sadosky, 2017.
- [8] S. Furber. *Shut down or restart? The way forward for computing in UK schools*. The Royal Society Education Section, 2012.
- [9] E. Grosclaude, R. Zurita, J. Riquelme, R. del Castillo, and M. Lechner. Designing A Myro-Compatible Robot For Education As Copyleft Hardware. In *CACIC 2014*, pages 372 – 382. UNLAM, 2014.
- [10] J. H. Maloney, K. Peppler, Y. Kafai, M. Resnick, and N. Rusk. Programming by choice: Urban youth learning programming with scratch. In *Proceedings of the 39th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '08*, pages 367–371, New York, NY, USA, 2008. ACM.
- [11] M. Prensky. Digital natives, digital immigrants part 1. *On the Horizon*, 9(5):1–6, 2001.
- [12] M. Resnick, J. Maloney, A. Monroy-Hernández, N. Rusk, E. Eastmond, K. Brennan, A. Millner, E. Rosenbaum, J. Silver, B. Silverman, and Y. Kafai. Scratch: Programming for all. *Commun. ACM*, 52(11):60–67, Nov. 2009.
- [13] J. Rodríguez, G. Grosso, R. Zurita, and L. Cecchi. Intervención de la facultad de informática en la enseñanza de ciencias de la computación en la escuela media basada en robótica educativa. In *XI Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2016)*, 2016.
- [14] J. Rodríguez and R. Zurita. *Proyecto de Extensión Vamos a la Escuela:Acercando las Ciencias de la Computación a la Escuela Media*. 2017, avalado por Resolución FaI 087/16.
- [15] M. Smith. Computer science for all. *The White House*, 2016.
- [16] C. Wilson, L. A. Sudol, C. Stephenson, and M. Stehlik. *Running on Empty: The Failure to Teach K–12 Computer Science in the Digital Age*. ACM and The Computer Science Teachers Association, 2010.
- [17] R. Zurita, J. de la Fuente, M. Bucarey, D. Bonet, R. del Castillo, J. Rodríguez, G. Grosso, and L. Cecchi. Mejorando las posibilidades de aprender a programar. ampliación del robot educativo multiplo n6 max a frankestito. In *XI Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2017)*, 2017.

Experiencias educativas, reflexiones, aportaciones docentes e investigaciones de la praxis en el contexto universitario

Mirta Fernandez; María V. Godoy; Sonia I. Mariño, Walter G. Barrios.

Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura.

9 de Julio 1449. CP: 3400. Corrientes. Argentina.

Universidad Nacional del Nordeste.

mirtagf@hotmail.com, mvvg2001@yahoo.com, simarinio@yahoo.com, waltergbarrios@yahoo.com

RESUMEN

En Educación Superior sus actores, como son los docentes y los investigadores, a través de sus actividades innovan al ser los generadores de ideas orientadas al aprendizaje activo.

En este sentido, surgen los diferentes grupos interdisciplinarios dirigidos a la aplicación crítica y reflexiva de métodos, herramientas y modelos a fin validar la educación y producir transformaciones. En el marco de un PI, se desarrollan y promueven diversas actividades de I+D+i vinculadas con las Tecnologías de la Información en las diversas modalidades que adopta la educación en el siglo XXI.

Palabras clave: aprendizajes activos, entornos educativos, procesos de aprendizajes.

CONTEXTO

El trabajo presenta los resultados de líneas de trabajo del proyecto “TI en los Sistemas de Información: Modelos, Métodos y Herramientas”, acreditado por la Secretaria General de la Universidad Nacional del Nordeste (UNNE). En este caso la indagación y aplicación de algunas líneas vinculadas con la enseñanza y aprendizajes en TIC.

Con relación a ello, mediante convenio con el Ministerio de Ciencia y Tecnología de la Nación, se desarrollaron las líneas investigativas en el marco del programa Piloto de Becas del Plan de Mejoramiento

de la Función I+D+i de la UNNE aprobada por Res. N° 995/14 C.S. UNNE.

El objetivo de la Convocatoria Piloto de Becas es contribuir a la formación de posgrado de alta calidad a nivel de maestría o doctorado de docentes de la UNNE, especialmente en áreas de vacancia o prioritarias, para que desarrollen en los próximos años conocimientos especializados en respuesta a demandas estratégicas de la región [1].

En este contexto se estudian las temáticas propuestas en el plan de trabajo denominado “Educación y NTICs: Entornos de Simulación como soporte a procesos de enseñanza - aprendizaje”, aprobado por Res. 250/15 C.S., (con vigencia 01/05/2015 al 31/04/2017). Se refutaron algunas hipótesis basadas en paradigmas y tendencia actuales en educación, se planteó la profundización en métodos y técnicas, que combinan la participación activa del estudiante y el acompañamiento del docente [2], [3], [4], [5], [6].

1. INTRODUCCIÓN

Siguiendo a Escudero [7], la innovación en los procesos de enseñanza-aprendizaje puede ser concebida como “*un determinado posicionamiento crítico y reflexivo que dirige sus esfuerzos tanto a validar la educación como a ir transformándola al servicio de valores debidamente legitimados ideológica,*

social, cultural, política y educativamente hablando”.

En los últimos años la innovación en la educación superior se ha visto acelerada en algunas instituciones que están transformando sus tradicionales métodos de enseñanza y aprendizaje, en los que los docentes y los investigadores son las principales fuentes de generación de ideas [8].

En una continua búsqueda de garantizar resultados de calidad en el aprendizaje para todos, a lo largo de toda la vida, como se propone en [9]; se hacen visibles las líneas investigativas con miras a producir dichas transformaciones.

La UNNE a través de sus diferentes proyectos acreditados, aporta sus mecanismos y esfuerzos con miras a fortalecer dicho proceso de transformación.

En relación a ello el PI propone como líneas investigativas en educación, el estudio de métodos y modelos y herramientas aplicables, a fin de alcanzar metas sustanciales en el aprendizaje de los alumnos que redunden en un aprendizaje significativo [10], [11].

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La tendencia de la integración de pedagogías para el aprendizaje más profundo en la educación superior ha ido adquiriendo cada vez más importancia desde hace algunos años [11]. Actualmente se desarrollan nuevas estrategias, especialmente en disciplinas relacionadas con las ciencias, las tecnologías, las ingenierías y las matemáticas conocidas como STEM (por sus siglas Science, Technology, Engineering y Mathematics) [8]. En relación en el PI mencionado y a través de la formación de sus RRHH, propone como líneas investigativas:

- i. *Estrategias y metodologías innovadoras de enseñanza.*
- ii. *Descripción y evaluación de experiencias innovadoras de enseñanza.*
- iii. *Herramientas estratégicas en el desarrollo de procesos educativos en la formación en tecnología aplicada.*
- iv. *Ubicuidad en los procesos de enseñar y aprender.*
- v. *El Aula virtual como mediadora del proceso de enseñanza y aprendizaje.*

Para ello, se indagaron métodos, modelos y herramientas TIC [8], [11], [12], [13]. Se vincularon saberes y experiencias, se evaluaron el impacto de las propuestas implementadas a fin de obtener discusiones en torno a los resultados.

3. RESULTADOS OBTENIDOS

Las experiencias se refieren a las configuraciones de las variables en la puesta en prácticas de los métodos, modelos y herramientas estudiadas.

En relación a las líneas identificadas como i) y ii) se indagó en la metodología de Aprendizaje Basado en Problemas (ABP) con la integración de materiales didácticos tridimensionales para la enseñanza en tecnologías aplicadas al Arte. Los resultados se plasman en un trabajo interdisciplinario y reflexivo que plantea formas alternativas de adquirir conocimientos a través del ABP [14].

Asimismo, se profundizó en el mecanismo que propone Aula Invertida como enfoque activo y centrado en el estudiante que surge de la premisa de extender el tiempo de una actividad con el objetivo de favorecer el pensamiento crítico y la autonomía en el aprendizaje; se revisó la literatura, dimensiones, implicancias empíricas y mediante una

experiencia se analizaron resultados obtenidos con miras a las mejoras en las implementaciones [15].

En lo que respecta a la línea iii) se indagó en el diseño y producción de software educativo de simulación, como un recurso interactivo para la enseñanza [16], [17].

Asimismo, tal como se mencionó, Se trabajó en la utilización de materiales en 3D como dispositivo mediador de la enseñanza y aprendizaje en el aula [14].

Otros resultados se plasmaron en [18], en el estudio de aspectos en la selección y uso de REA (Recursos Educativos Abiertos) en la implementación de los lineamientos de Aula Invertida.

Vinculadas a las líneas de trabajo iv) y v) se estudiaron consideraciones en el diseño de aula virtual como apoyo a la estrategia de Aula Invertida [19].

Desde un punto de vista holístico y de la gestión, se exploró en aspectos de la expansión de la educación virtual y se reseñaron las estrategias abordadas desde del programa UNNE Virtual [20].

4. FORMACIÓN DE RECURSOS HUMANOS

En el marco de la beca del programa Piloto de Becas del Plan de Mejoramiento de la Función I+D+i de la UNNE, la becaria e integrante del PI, Lic. Fernandez Mirta promovió la totalidad de los cursos y seminarios de la Maestría en Tecnología Aplicada a la Educación que ofrece la Universidad de La Plata (UNLP), el cual ofrece conocimiento actualizado vinculado a las metodologías, tecnologías y herramientas que brinda la Ciencia Informática, y que cobran sentido en el marco de procesos educativos, de manera tal de favorecerlos y fortalecerlos.

Se realizó el estudio de teorías, métodos y técnicas, diversas y muy enriquecedoras para el quehacer docente.

Asimismo, en 2015 la becaria accedió a la categoría V del programa de incentivo de docentes investigadores.

Las tareas desarrolladas estuvieron coordinadas y supervisadas por la directora del plan de beca y co-directora del PI, Prof. Godoy Guglielmone, María Viviana y la directora del PI, Prof. Mariño Sonia I. Colaboró en los estudios y las producciones el integrante Lic. Barrios Walter.

5. BIBLIOGRAFIA

- [1] RESOLUCIÓN N° 995/14 C.S. UNNE. Reglamento Becas Piloto en Áreas de Vacancia y/o Prioritarias. Disponible en: http://www.unne.edu.ar/component/joomla/doc/Investigacion/Regl_Becas_VacanciaPrioritarias.pdf/download
- [2] Adell, J., & Castañeda, L. (2012). Tecnologías emergentes, ¿pedagogías emergentes. Tendencias emergentes en educación con TIC 13-32.
- [3] Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)? Contemporary issues in technology and teacher education, 9(1), 60-70.
- [4] Harris, J., Mishra, P., & Koehler, M. (2009). Teachers' technological pedagogical content knowledge and learning activity types: Curriculum-based technology integration reframed. *Journal of Research on Technology in Education*, 41(4), 393-416.
- [5] Cabero, J.: La formación del profesorado en TIC: modelo TPACK, Ed. Universidad de Sevilla (2014)
- [6] Cabero Almenara, J. (2015). Reflexiones educativas sobre las tecnologías de la información y la comunicación (TIC). *Tecnología, Ciencia y Educación* 1, 19-27.

- [7] Escudero, J. M. (1995). La integración de las nuevas tecnologías en el currículum y en el sistema escolar, en Rodríguez Diéguez, J.L. y Sáenz, o. (Coords.): *Tecnología Educativa. Nuevas Tecnologías aplicadas a la Educación*. Alcoy, Marfil; 397-412.
- [8] Becker, S. A., Cummins, M., Davis, A., Freeman, A., Hall, C. G., & Ananthanarayanan, V. (2017). *NMC Horizon Report: 2017 Higher Education Edition* (pp. 1-60). The New Media Consortium.
- [9] Amadio, M., Opertti, R., & Tedesco, J. C. (2015). El currículo en los debates y en las reformas educativas al horizonte 2030: Para una agenda curricular del siglo XXI.
- [10] Ausubel, D. P., Novak, J. Y. H. H., & Hanesian, H. (1976). Significado y aprendizaje significativo. *Psicología educativa: un punto de vista cognoscitivo*, 53-106.
- [11] Acuña, B. P. (2015). *Vectores de la pedagogía docente actual*. ACCI (Asoc. Cultural y Científica Iberoameric.).
- [12] Ortega, J. H., Fruscio, M. P., López, D. S., & Gutiérrez, A. V. (2012). Tendencias emergentes en educación con TIC. *Asociación Espiral Educación y Tecnología*, 1, 286.
- [13] Castaño, C., & Cabero, J. (2013). *Enseñar y aprender en entornos m-learning*. Madrid: Síntesis.
- [14] Fernández, M. G., & Godoy Guglielmone, M. V. (2016). Integración de materiales didácticos en 3D para el aprendizaje basado en problemas en la enseñanza de artes. In XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016).
- [15] Fernández, M. G., G. Guglielmone, M. V., Mariño, S. I., & Barrios, W. G. (2017, October). Inverting the Class or Investing in the Class? Flipped Classroom in Teaching Technology for Multimedia Production. In Argentine Congress of Computer Science (pp. 295-307). Springer, Cham.
- [16] Fernández, Mirta G.; Godoy, Maria V.; Mariño Sonia I. (2016): Fernández, M. (2016). Desarrollo de software educativo hipermedia basado en Modelo Instruccional. 7mo. Seminario Internacional de Educación a Distancia. Univ. Nacional del Litoral.
- [17] Fernández, M. G., Godoy, M. V. and Mariño, S. (2016). Diseño de material multimedia para enseñanza aprendizaje. XXII Reunión de Comunicaciones Científicas y Tecnológicas. UNNE. Corrientes.
- [18] Fernández, M, & Godoy M. V.. (2017). Diseño de espacio virtual para la utilización de Aula Invertida. 4º Jornadas de TIC e Innovación en el Aula – Univ. Nac. de La Plata.
- [19] Castro Chans, N., Oviedo L., Godoy Guglielmone, M. V, Fernández M. (2017). El programa UNNE Virtual, su desarrollo y prospectiva en el marco de la gestión académica. VIII Encuentro Nacional y V Latinoamericano, La Universidad como objeto de investigación.

Entorno Web para la Programación del Robot Frankestito

Marcos Picucci Jorge Rodríguez Rafael Zurita Laura Cecchi

email: marcos.picucci@est.fi.uncoma.edu.ar {j.rodrig.rafa, lcecchi}@fi.uncoma.edu.ar,

Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

Resumen

La robótica educativa ha emergido en la última década en el contexto nacional como una herramienta adecuada para la enseñanza de las Ciencias de la Computación en el ámbito de la educación obligatoria, en múltiples niveles del sistema educativo.

En este trabajo se presenta una Línea de Investigación que propone el desarrollo de una arquitectura cliente-servidor basada en Web para la programación del robot educativo *Frankestito*.

El sistema propuesto proveerá una interfaz web que permita a los usuarios codificar, guardar, cargar, compilar y visualizar los resultados de sus programas, de modo que no se requiera instalar ningún software en sus computadoras. El ambiente permitirá programar en los diferentes lenguajes de programación soportados actualmente por *Frankestito*.

La interfaz web dará la posibilidad de ejecutar los programas sobre un simulador o el robot físico. En este último caso, se espera que el usuario no solo pueda observar el comportamiento del robot sino también un streaming de lo que la cámara del robot captura.

Palabras Clave: ROBÓTICA EDUCATIVA, EDUCACIÓN CIENCIAS DE LA COMPUTACIÓN, ENSEÑANZA DE LA PROGRAMACIÓN, ENTORNO WEB PARA ROBÓTICA EDUCATIVA, PROGRAMACIÓN ON-LINE DE ROBOTS. LABORATORIO

REMOTO.

Contexto

Esta línea de investigación se desarrolla, por un lado, en el contexto de los temas de interés que promueve el Grupo de Investigación en Lenguajes e Inteligencia Artificial (GILIA), de la Facultad de Informática. En particular, se enmarca en el ámbito de dos proyectos de investigación miembros del GILIA, ambos financiados por la Universidad Nacional del Comahue y con una duración de cuatro años a partir de enero del 2017: *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación (04/F015)* y *Agentes Inteligentes y Web Semántica (04/F014)*.

Por otro lado, el trabajo está en el contexto del Convenio Marco de Colaboración firmado durante 2016 entre la Facultad de Informática y el Ministerio de Educación de la Provincia del Neuquén. Particularmente, se trabaja con el Consejo Provincial de Educación de la Provincia de Neuquén.

Las actividades concretadas en forma conjunta entre la Facultad de Informática y el Consejo Provincial de Educación, en el ámbito de la investigación se plantean como articuladas a un conjunto de proyectos de extensión [4, 12], en ejecución durante 2017 y 2018, con intención de construir y ampliar conocimiento a partir de la revisión y análisis de resultados desarrollados en el campo de la praxis.

1. Introducción

La robótica en el ámbito educacional se ha puesto al descubierto como una herramienta con el potencial necesario para afrontar la para nada trivial tarea de enseñar y aprender conceptos de las Ciencias de la Computación en múltiples niveles del sistema educativo. Mediante una eficaz puesta en práctica de este enfoque didáctico se puede promover el trabajo colaborativo, generar interés sobre diversas disciplinas académicas e impulsar el pensamiento computacional, crítico y creativo [5, 3, 1, 9, 2].

En el ámbito de la Facultad de Informática, se ha desarrollado a *Frankestito*, un robot educativo de bajo costo distribuido bajo licencias Open Source y Open Hardware. Cuenta con capacidad de visión, movimiento y comunicación vía WiFi [7].

En línea con esta iniciativa y en el mismo ámbito de trabajo, se consiguió extender posteriormente a los robots educativos *Multiplo N6 Max* para que cuenten con las mismas características funcionales que *Frankestito* [13].

Desde 2013, la Facultad de Informática desarrolla proyectos de extensión para promover la enseñanza de las Ciencias de la Computación en la Escuela Secundaria utilizando a la robótica educativa, en particular robots compatibles con *Frankestito*, como eje motivador [11, 12, 4].

En estas actividades, en lo que respecta a la utilización del robot por parte de los estudiantes, se debe realizar un trabajo previo de instalación y configuración de software que, ineludiblemente, se traduce en una contrariedad a resolver por el docente a cargo del curso. En general el docente cuenta con tiempos limitados y no siempre posee los conocimientos técnicos necesarios para solucionar problemas de instalación específicos, como los casos en los que debe resolver dependencias de software o actualización de bibliotecas.

Tomando en consideración la experiencia obtenida dentro de las aulas, se infiere la necesidad de una mejora sobre el modelo de interacción actual entre docentes, estudiantes,

computadoras y robots.

En esta Línea de Investigación se propone el desarrollo de un ambiente web para la programación de robots educativos compatibles con *Frankestito*, de forma de independizar y aislar su comunicación con los usuarios en un único servidor. Dicho entorno cliente-servidor posibilitará programar a *Frankestito* en línea, permitiendo ejecutar el programa sobre un simulador o un robot físico accediendo a un laboratorio remoto.

Si bien existen herramientas con objetivos similares a los que en este trabajo se plantean [8, 10, 6], cada una trabaja sobre contextos de hardware y software específicos y por lo tanto ajenas a las características de *Frankestito*.

El trabajo presentado está estructurado como sigue. En la siguiente sección se introducen los proyectos de investigación que dan contexto al trabajo y la línea en desarrollo. En la sección 3 se detallan los avances en este trabajo y los trabajos futuros. Finalmente, se comentan aspectos en relación a la formación de recursos humanos.

2. Líneas de investigación y desarrollo

Esta línea de investigación que propone el desarrollo de un entorno web para la programación del robot educativo *Frankestito* surge en el contexto de dos proyectos de investigación.

Por una parte, el proyecto de investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación*, busca desarrollar modelos que contribuyan a la producción de un marco teórico, a fin de asistir la inclusión de la computación en la educación. Alineado con el modelo subyacente, se plantea como objetivo el diseño e implementación de herramientas que los soporten.

El desarrollo de arquitecturas que permitan la programación de robots para generar comportamiento inteligente es un objetivo específico del proyecto de investigación *Agentes Inteligentes y Web Semántica*, que tiene

como objetivo generar conocimiento especializado en el área de agentes inteligentes.

Así, el desarrollo de una arquitectura cliente-servidor basada en tecnologías Web que permita a los usuarios de la comunidad educativa programar en línea a *Frankestito* es de interés para ambos proyectos, colaborando en su diseño e implementación. Los resultados de este trabajo permitirán ampliar las posibilidades didácticas de *Frankestito* y facilitar la enseñanza y aprendizaje de la programación.

Como parte de esta propuesta, se desea que la aplicación Web cuente con las siguientes características funcionales:

- *Selección del lenguaje de programación:* Los usuarios podrán seleccionar el lenguaje de programación de su preferencia, entre los soportados por *Frankestito*, para codificar sus programas. Actualmente, se puede interactuar con el robot a través de una diversidad de lenguajes de programación como Python, C y C++, entre otros. La arquitectura diseñada deberá ser suficientemente robusta como para ser expandida a otros lenguajes de programación como Scratch y Prolog, entre otros.
- *Editor de Código en línea:* De esta forma se permitirá a los estudiantes codificar y compilar los programas que el robot posteriormente ejecutará. Se pretende que el editor cuente con algún tipo de asistencia al usuario como la detección de errores sintácticos.
- *Gestión de usuarios y archivos:* Por un lado, se deberá diferenciar a los diferentes tipos de usuarios con sus respectivos privilegios. Por ejemplo, estudiante, docente y administrador.
Por el otro lado, se deberá permitir a los estudiantes almacenar programas en su perfil local y cargar programas externos al mismo.
- *Ejecución del programa sobre un simulador o un robot físico:* El entorno Web

deberá proveer acceso al simulador y al robot físico para visualizar el comportamiento del programa. El usuario será quién seleccione dónde quiere ejecutar su programa.

- *Laboratorio remoto para el acceso al robot físico:* El entorno deberá proveer acceso a un laboratorio remoto y una ventana para mostrar en vivo los movimientos. Asimismo, se desea que el usuario pueda ver un streaming de lo que está capturando la cámara del robot. A través de esta funcionalidad se ayudará en la tarea de debugging de los programas.
- *Gestión de turnos para el laboratorio remoto:* Para evitar que un usuario se apropie indefinidamente del robot físico en el laboratorio remoto, se deberá contar con un sistema de turnos de corta duración que de lugar el uso equitativo del recurso en cuestión.

3. Resultados obtenidos y esperados

En una primer etapa, se analizó la arquitectura de la plataforma *Frankestito* y la adaptación del robot *Multiplo N6 Max* [13], compatible con el primero mencionado, y se realizaron trabajos de campo para una mejor comprensión del comportamiento del robot. Se estudiaron principalmente las funcionalidades de comunicación WiFi.

En paralelo, se comenzó con una revisión sobre diferentes ambientes Web que permiten programar en forma on-line robots educativos, a fin de identificar sus características principales.

Actualmente, se está trabajando en el diseño de la arquitectura cliente-servidor basada en Web.

La arquitectura cliente-servidor resultante de esta herramienta deberá ser modular

y robusta como para soportar todas las características mencionadas en la sección 2, como también posibles extensiones. Asimismo, se espera que pueda soportar la posibilidad de evolucionar hacia una herramienta colaborativa que facilite y promueva el trabajo en equipo de un grupo de estudiantes.

Como resultado de la implementación de esta herramienta se espera simplificar y enriquecer la interacción entre usuarios y robots. Bajo esta propuesta, solo se deberá requerir de una computadora o dispositivo móvil con un navegador web. En este sentido, se busca dejar de lado el procedimiento actual, específico para cada computadora, de instalación y configuración de software .

Por otra parte, se espera que de manera indirecta se produzca un aplanamiento de la curva de aprendizaje. En forma complementaria, el proceso podrá desarrollarse fuera del ámbito institucional, específicamente fuera de los tiempos estrictos de uso de laboratorios. De esta manera se busca que los estudiantes pongan en práctica con mayor frecuencia las fases de codificación, ejecución y corrección que afectan en forma directa y positiva en el proceso de aprendizaje [5].

Por último, se desea que el trabajo realizado pueda ser tomado como base para futuros trabajos de investigación, es decir que sea susceptible a ser extendido o modificado para implementar nuevas funcionalidades sin la necesidad de realizar cambios complejos sobre la arquitectura general del sistema.

4. Formación de Recursos Humanos

Uno de los autores de este trabajo se encuentra desarrollando su tesis de grado de la Licenciatura en Ciencias de la Computación, en la temática de esta línea de investigación.

Por otra parte, otro de los autores de este trabajo está inscripto en la Maestría en Enseñanza en Escenarios Digitales que desarrollan de manera conjunta las Universidades Nacionales de Cuyo, Comahue, Patagonia Austral, Patagonia San Juan Bosco, San

Luis, Chilecito y La Pampa.

Referencias

- [1] S. Atmatzidou and S. Demetriadis. Advancing students' computational thinking skills through educational robotics: A study on age and gender relevant differences. *Robotics and Autonomous Systems*, 75:661–670, 2016.
- [2] R. D. Beer, H. J. Chiel, and R. F. Drushel. Using autonomous robotics to teach science and engineering. *Communications of the ACM*, 42(6):85–92, 1999.
- [3] F. B. V. Benitti. Exploring the educational potential of robotics in schools: A systematic review. *Computers & Education*, 58(3):978–988, 2012.
- [4] L. Cecchi and G. Grosso. *Proyecto de Extensión Agentes Robots: Divulgando Computación en la Escuela Media*. 2017, avalado por Resolución FaI 088/16.
- [5] B. Fagin and L. Merkle. Measuring the effectiveness of robots in teaching computer science. In *ACM SIGCSE Bulletin*, volume 35, pages 307–311. ACM, 2003.
- [6] J. Fernandez and A. Casals. Open laboratory for robotics education. In *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, volume 2, pages 1837–1842. IEEE, 2004.
- [7] E. Grosclaude, R. Zurita, R. d. Castillo, M. Lechner, and J. Riquelme. Designing a myro-compatible robot for education as copyleft hardware. In *XX Congreso Argentino de Ciencias de la Computación (Buenos Aires, 2014)*, 2014.
- [8] B. Jost, M. Ketterl, R. Budde, and T. Leimbach. Graphical programming environments for educational robots:

- Open roberta-yet another one? In *Multi-media (ISM), 2014 IEEE International Symposium on*, pages 381–386. IEEE, 2014.
- [9] K. J. O’Hara and J. S. Kay. Investigating open source software and educational robotics. *Journal of Computing Sciences in Colleges*, 18(3):8–16, 2003.
- [10] P. Petrovi and R. Balogh. Deployment of remotely-accessible robotics laboratory. *International Journal of Online Engineering (iJOE)*, 8(S2):31–35, 2012.
- [11] J. Rodríguez, G. Grosso, R. Zurita, and L. Cecchi. Intervención de la facultad de informática en la enseñanza de ciencias de la computación en la escuela media basada en robótica educativa. In *XI Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2016)*, 2016.
- [12] J. Rodríguez and R. Zurita. *Proyecto de Extensión Vamos a la Escuela:Acercando las Ciencias de la Computación a la Escuela Media*. 2017, avalado por Resolución FaI 087/16.
- [13] R. Zurita, J. d. I. Fuente, M. Bucarey, D. Bonet, R. d. Castillo, G. Grosso, L. Cecchi, J. Rodríguez, et al. Mejorando las posibilidades de aprender a programar, ampliación del robot educativo multiplo n6 max a frankstato. In *XII Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET, La Matanza 2017)*., 2017.

El lugar de las Ciencias de la Computación en el currículum de la Escuela Secundaria Argentina

Sonia Sommer² María Eda Cornejo² Marcos Cortez²
Jorge Rodríguez¹

soniasommer@yahoo.com, medacornejo@gmail.com, cortezmarcos@gmail.com,
j.rodri@fi.uncoma.edu.ar

¹*Grupo de Investigación en Lenguajes e Inteligencia Artificial*
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

²*Consejo Provincial de Educación*
MINISTERIO DE EDUCACIÓN DE LA PROVINCIA DE NEUQUÉN

Resumen

La enseñanza de conceptos de Ciencias de la Computación en la escuela es reconocida como prioritaria y ha logrado altos niveles de consenso.

En este sentido, varias Universidades Nacionales, programas nacionales como Program.ar y PLANIED y la Fundación Sadosky han articulando acciones orientadas a mejorar la aproximación de las Ciencias de la Computación a la Escuela Secundaria.

En Argentina, la Ley de Educación Nacional establece que la revisión de la estructura curricular de la Educación Secundaria corresponde a las distintas jurisdicciones. Por lo tanto, el diseño curricular de cada provincia asigna a la computación lugares y roles específicos.

En este trabajo se presenta una Línea de Investigación que busca describir rigurosamente el panorama de la computación en los diseños curriculares vigentes para la Escuela Secundaria en Argentina. Se propone analizar la posición que ocupa en las propuestas formativas y aportar elementos que colaboren con los procesos de discusión tendientes

a construir nuevos diseños.

Palabras Clave: CIENCIAS DE LA COMPUTACIÓN, CURRÍCULUM, ESCUELA SECUNDARIA, ENSEÑANZA DE LA COMPUTACIÓN, REVISIÓN SISTEMÁTICA.

Contexto

Esta propuesta se ubica en el contexto de las iniciativas promovidas por el Grupo de Investigación en Lenguajes e Inteligencia Artificial de la Facultad de Informática y del Convenio Marco de Colaboración firmado durante 2016 entre la Facultad de Informática y el Ministerio de Educación de la Provincia del Neuquén.

Este trabajo se desarrolla en el ámbito del proyecto de investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación (04/F015)* que está financiado por la Universidad Nacional del Comahue a través de la Secretaría de Ciencia y Técnica y por el Consejo Provincial de Educación en el contexto del Convenio Marco de Colaboración. El proyecto tiene prevista una duración de cuatro años a partir de enero del 2017.

1. Introducción

La enseñanza de conceptos de Ciencias de la Computación en la escuela es reconocida como prioritaria y cuenta con gran consenso, tanto en países desarrollados como en desarrollo. Esta tendencia involucra a gobiernos, organizaciones civiles, grupos de investigación y docentes [9, 13].

En este contexto se considera a la computación tan relevante como otras disciplinas históricamente ponderadas, tales como matemática, historia o lengua y literatura. Esto se debe a que resulta necesaria para mejorar las posibilidades de entender e intervenir el mundo que rodea a los estudiantes.

En este sentido, en Argentina, varias Universidades Nacionales, programas nacionales como Program.ar y PLANIED y la Fundación Sadosky han articulado acciones que proponen la implementación de estrategias orientadas a aproximar las Ciencias de la Computación a la Escuela Secundaria.

En este marco, en 2015 el Consejo Federal de Educación declara de importancia estratégica la enseñanza y el aprendizaje de la programación en todas las escuelas durante la escolaridad obligatoria.

A partir de 2005, la Facultad de Informática de la Universidad Nacional del Comahue que abarca las provincias de Río Negro y Neuquén, establece vínculos de colaboración con varias escuelas del nivel medio de la región con la intención de promover la inclusión progresiva y sostenida, en las propuestas de enseñanza, de contenidos relacionados a las Ciencias de la Computación[15, 14].

Sin embargo, más allá de los esfuerzos realizados y los acuerdos construidos, varios reportes muestran que la incorporación de forma rigurosa y de manera sostenible de las Ciencias de la Computación en la educación secundaria es un proceso en desarrollo en Argentina, así como en la mayoría de los países[6, 5, 16].

En Argentina, la Ley de Educación Nacional establece que la revisión de la estructura curricular de la Educación Secundaria corresponde a las distintas jurisdicciones [11].

Por lo tanto cada diseño curricular tiene carácter singular e independiente y asigna a la computación lugares y roles específicos. Esta situación produce un alto grado de dispersión en relación a los enfoques y perspectivas con que la computación se integra en las propuestas formativas.

El contexto descripto evidencia la necesidad de desarrollar líneas de investigación, específicas al campo de la Educación en Ciencias de la Computación, que contribuyan a describir y comprender la situación de la enseñanza de la disciplina en el país.

En este trabajo se presenta una Línea de Investigación que busca describir rigurosamente el panorama de la computación en los diseños curriculares vigentes para la Escuela Secundaria en Argentina. Se propone analizar la posición que ocupa la computación en las propuestas formativas y aportar elementos que colaboren en los procesos de discusión tendientes a construir nuevos diseños.

2. Línea de Investigación

El estudio a desarrollar en esta Línea de Investigación consiste en revisar y analizar el lugar que ocupa la computación en los diseños curriculares vigentes para la Escuela Secundaria en Argentina.

En esta sección identificamos las características a observar durante el proceso de revisión y proponemos una estrategia metodológica, que toma como principal referencia los aportes elaborados para realizar revisiones sistemáticas de literatura[10, 2], para explorarlas.

Con intención de favorecer la comprensión sobre las perspectivas expresadas en las propuestas curriculares, el proceso de revisión necesita de la definición de una serie de aspectos a explorar:

- Rol y propósito de la computación en la escuela

Considerar diferentes roles y propósitos para la computación en la escuela y en

los planes de estudio conduce a la definición de opciones distintas en relación a los lugares que se asigna a la disciplina en el diseño del currículum[7].

Las computadoras se incorporan a la escuela con diferentes roles, tales como la mejora de la calidad de la educación, la alfabetización digital de la población, el desarrollo de competencias TIC, la reducción de la brecha digital que se observa en las sociedades y la construcción de habilidades para la resolución de problemas en el marco del pensamiento computacional [3].

Reportes y estudios realizados recientemente sobre las propuestas curriculares, principalmente en Europa y Estados Unidos, identifican además la necesidad de reconocer a las Ciencias de la Computación como disciplina académica rigurosa e impulsan el desarrollo de iniciativas en esta dirección[9, 4, 17].

- Posición de la computación en el plan de estudio

Varios estudios y reportes plantean que un aspecto central a considerar en los procesos de revisión relacionados a la computación en los diseños curriculares para la educación obligatoria es la posición asignada a la disciplina en los planes de estudio[6, 12].

En este marco, el debate está planteado entre posicionar a la computación como disciplina académica independiente y los enfoques que proponen integrar algunas estrategias y recursos del pensamiento computacional en otros espacios curriculares[16].

Los argumentos en favor de considerarla una disciplina escolar independiente han constituido una tendencia que impulsó procesos de reforma curricular en varios países [1, 16].

Además, se presta atención a los momentos en los que la disciplina aparece en los planes de estudio y a la carga horaria asignada al área.

- Enfoques, perspectivas y selección de contenidos

En el reporte *Shut down or restart?* se establecen algunas definiciones de trabajo que proponen considerar la posibilidad de desagregar en áreas, claramente definidas, como *Alfabetización Digital*, *Tecnología de la Información y Ciencias de la Computación*[4].

Esta propuesta, con algunos ajustes, constituye la base para las definiciones adoptadas en este trabajo. A la desagregación elaborada por la *Royal Society* sumamos la categoría de análisis *Mejoramiento de los Aprendizajes*.

Alfabetización Digital, plantea el desarrollo de competencias digitales básicas. Es decir el conjunto de habilidades para usar satisfactoriamente las TIC [4, 3].

Competencias TIC, considera la construcción y empleo de estrategias para utilizar sistemas informáticos preexistentes para satisfacer necesidades relacionadas a campos específicos como la industria, el comercio o el arte [4, 8].

Ciencias de la Computación, propone considerar una disciplina académica rigurosa que abarca conceptos y prácticas computacionales fundamentales[4, 9].

Mejoramiento de la Calidad de los Aprendizajes, contempla la introducción transversal en el proceso de enseñanza con el propósito de mejorar los logros educativos[3, 8].

Las características expuestas no se expresan en forma aislada en los planes de estudio, se considera que presentan una relación dialéctica que las define mutuamente y en conjunto describen el lugar asignado a la computación en la propuesta formativa.

La línea de investigación busca identificar roles y propósitos de la computación en la Escuela Secundaria Argentina, la posición que ocupa en los diseños curriculares vigentes y

los enfoques expresados en las propuesta formativas.

2.1. Metodología

Para favorecer la comprensión de las perspectivas expresadas en las propuestas curriculares se considera necesario avanzar en la elaboración de una revisión sistemática y rigurosa del panorama de la computación en el currículum escolar.

La metodología propuesta para elaborar la descripción es una revisión sistemática que contempla a los documentos curriculares como insumo primario. Este es un proceso desarrollado para identificar, evaluar e interpretar la información destacada de una colección de literatura de interés para la investigación, realizando la búsqueda y extracción de lo más relevante de acuerdo a criterios que se definen explícitamente[2, 10].

Las etapas que componen la metodología son[10]:

- Planificar la revisión

Esta primer etapa consiste en establecer el protocolo de revisión en el que se definen los antecedentes, la estrategia y los términos de búsqueda, los criterios de análisis y selección para la extracción de datos, así como la agenda del proyecto.

- Realizar la revisión

Esta etapa es determinada por la revisión de la literatura y contempla la estrategia de búsqueda, los criterios de selección que se han determinado para la elección de los estudios primarios, la elaboración de formularios para la extracción de datos, el análisis de la información de forma cuantitativa y/o cualitativa finalizando con la síntesis de los datos.

- Resultados de la revisión

Los resultados y conclusiones de la investigación se presentan en una nueva publicación académica.

3. Resultados

Inicialmente, se realizó una recopilación de los diseños curriculares vigentes para la Escuela Secundaria en Argentina. De la misma resultó una compilación de 19 propuestas curriculares. Es importante mencionar que las provincias restantes están actualmente en proceso de reforma curricular.

Se llevó a cabo un estudio sobre diferentes estrategias para avanzar en el proceso de revisión sistemática. Del mismo surgió la estrategia metodológica planteada en la sección anterior.

Se realizó un análisis sobre trabajos que estudian la situación de las Ciencias de la Computación en propuestas curriculares en Estados Unidos y Europa.

En el contexto de esta Línea de Investigación se espera obtener los siguientes resultados:

- Presentar una síntesis que contribuya a describir y comprender la situación de la enseñanza de las Ciencias de la Computación en el país.
- Avanzar en la identificación de objetivos, paradigmas, enfoques y perspectivas expresados en las propuestas curriculares vigentes para la Escuela Secundaria en Argentina.
- Aportar elementos que colaboren con el análisis de la situación actual y la definición de perspectivas futuras.

4. Formación de Recursos Humanos

Se espera que el desarrollo de esta Línea de Investigación contribuya a la formación de recursos humanos en el campo de las Ciencias de la Computación en la Educación.

En este sentido tres de los autores de este artículo cursan maestrías orientadas a conocer, comprender y analizar procesos relacionados con las tecnologías en la educación.

Referencias

- [1] S. Bocconi, A. Chiocciariello, G. Dettori, A. Ferrari, K. Engelhardt, P. Kamylyis, and Y. Punie. Developing computational thinking in compulsory education. *European Commission, JRC Science for Policy Report*, 2016.
- [2] A. Booth, A. Sutton, and D. Papaioannou. *Systematic approaches to a successful literature review*. Sage, 2016.
- [3] CEPAL, OEI, and S. G. Iberoamericana. *Metas educativas 2021*. Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI), 2010.
- [4] S. Furber. *Shut down or restart? The way forward for computing in UK schools*. The Royal Society Education Section, 2012.
- [5] Google and Gallup. *Searching for computer science: Access and barriers in U.S. K-12 education*. 2015.
- [6] Google and Gallup. Trends in the state of computer science in u.s. k-12 schools. 2016.
- [7] M. Guzdial. Learner-centered design of computing education: Research on computing for everyone. *Synthesis Lectures on Human-Centered Informatics*, 8(6):1-165, 2015.
- [8] ISTE. ISTE standards for students". *ISTE Standards*, 2016.
- [9] K-12 Computer Science Framework Steering Committee. *The K-12 Computer Science Framework*. ACM, 2016.
- [10] B. Kitchenham and S. Charters. Guidelines for performing systematic literature reviews in software engineering. EBSE technical report. *Keele University & Department of Computer Science University of Durham*, 2007.
- [11] Ley Nacional. 26.206 Ley de Educación Nacional. *Disposiciones Generales*, 26 y 27 de Abril de 2018 2006.
- [12] N. Reynolds, D. P. Chambers, M. M. Syslo, A. Fluck, M. Cox, C. Angeli, J. Malyn-Smith, J. Voogt, J. Zagami, P. Micheuz11, et al. Computer science in the school curriculum: Issues and challenges. In *Tomorrow's Learning: Involving Everyone. Learning with and about Technologies and Computing: 11th IFIP TC 3 World Conference on Computers in Education, WCCE 2017, Dublin, Ireland, July 3-6, 2017, Revised Selected Papers*, volume 515, page 421. Springer, 2018.
- [13] F. Sadosky. *CC - 2016 Una propuesta para refundar la enseñanza de la computación en las escuelas Argentinas*. Fundación Sadosky, Argentina, 2013.
- [14] S. Sommer, M. E. Cornejo, G. Grosso, and J. Rodríguez. Construyendo aplicaciones móviles en la escuela: un enfoque para la enseñanza y el aprendizaje de la programación. In *XII Congreso de Tecnología en Educación y Educación en Tecnología*, 2017.
- [15] S. Sommer, J. E. Sznek, and J. Rodríguez. Divulgando temáticas computacionales-internet segura. In *X Congreso sobre Tecnología en Educación & Educación en Tecnología*, 2015.
- [16] M. Webb, N. Davis, T. Bell, Y. J. Katz, N. Reynolds, D. P. Chambers, and M. M. Syslo. Computer science in k-12 school curricula of the 21st century: Why, what and when? *Education and Information Technologies*, 22(2):445-468, 2017.
- [17] C. Wilson, L. A. Sudol, C. Stephenson, and M. Stehlik. *Running on Empty: The Failure to Teach K-12 Computer Science in the Digital Age*. ACM and The Computer Science Teachers Association, 2010.

La Universidad y los procesos de transformación: el Género en las TIC. El caso de la Facultad de Informática de la UNLP

Lía Molinari, Claudia Queiruga, Ana Ungaro, Sofía Martín, Ana Paola Amadeo, María Alejandra Osorio, Javier Díaz

Laboratorio de Investigación de Nuevas Tecnologías Informáticas (LINTI). Facultad de Informática. Universidad Nacional de La Plata
50 y 120. La Plata

{lmolinari, claudiaq, anaungaro, pamadeo, javierd}@info.unlp.edu.ar,
aosorio@cespi.unlp.edu.ar, smartin@linti.unlp.edu.ar

RESUMEN

La brecha digital de género persiste y se agrava a pesar que los usos y accesos a las TIC son cada vez mayores. Actualmente el problema de la brecha digital de género se identifica como el problema de las mujeres en las TIC, en el que cada vez es más evidente la inferioridad numérica de las mujeres en el ámbito de los estudios, investigación y la profesión del sector TIC, en especial en los países de Occidente. Esta escasa participación de las mujeres es especialmente visible en la formación universitaria, en la que existe un evidente sesgo de género en relación a la elección de las carreras, que tiene como resultado que las mujeres son claramente una minoría en las carreras tecnológicas, y son especialmente preocupantes los datos relativos a las carreras de Informática. La Facultad de Informática de la UNLP no es ajena a este problema, es por ello que esta línea de investigación del Laboratorio de Investigación en Nuevas Tecnologías Informáticas (LINTI) propone trabajar en la búsqueda de las causas concretas y las posibles soluciones de la brecha digital de género.

Palabras clave: género, brecha digital de género, TIC, Informática, STEM

CONTEXTO

La línea de investigación "Género en las TIC" presentada en este trabajo, está inserta en el proyecto de investigación "Internet del Futuro: Ciudades Digitales Inclusivas, Innovadoras y Sustentables, IoT, Ciberseguridad, Espacios de Aprendizaje del Futuro" del Programa Nacional de Incentivos a docentes-investigadores, que se desarrolla en el Laboratorio de Investigación en Nuevas Tecnologías Informáticas (LINTI) de la Facultad de Informática de la Universidad Nacional de La Plata (UNLP). Este proyecto

está acreditado por la UNLP y financiado por partidas del presupuesto nacional. Asimismo, esta línea de trabajo articula los procesos de investigación con proyectos de extensión de la UNLP, de los que participan los docentes-investigadores del LINTI, dándole el sustento de pertinencia local y regional a los resultados obtenidos.

1. INTRODUCCIÓN

Actualmente podemos afirmar que la primera brecha digital de género, vinculada al acceso desigual de las mujeres a las TIC y que explica su escasa participación en las mismas (Gürer, D. et al, 2002), se ha ido superando y hoy las mujeres son usuarias de TIC en cifras similares a los hombres, inclusive los superan en tecnologías como celulares y redes sociales. Una vez superada esta primera brecha, se observa una segunda brecha digital (Castaño C., 2008, 2009) que remite a la desigualdad en cuanto a la intensidad, los usos, las habilidades y el apego respecto de las TIC y, es esta segunda brecha la que explica la relación desigual entre hombres y mujeres respecto de las TIC y, como consecuencia de ello, la escasa participación de las mujeres en relación al uso experto de las tecnologías digitales. Este uso experto está estrechamente vinculado al acceso a los estudios universitarios, dado que solamente aquellas mujeres que tengan interés en desarrollar prácticas expertas, accederán a la universidad. De acuerdo a lo señalado por Cecilia Castaño, la brecha digital de género "está relacionada con el dominio masculino de la áreas estratégicas de la educación, la investigación y el empleo relacionado con las ciencias, las ingenierías y las TIC" (Castaño C., 2008). La escasa participación de las mujeres en las TIC se manifiesta de manera casi generalizada en las matriculaciones en las universidades, en donde se observa que cada vez hay más mujeres estudiando en todo el mundo, inclusive es mayor en países de la

Unión Europea, Estados Unidos y en algunos países de América Latina (UNESCO, 2013) (UNESCO, 2005). Sin embargo existe un importante sesgo en cuanto a la elección de las carreras, siendo las mujeres una minoría en las carreras tecnológicas y dentro de las mismas son preocupantes los datos relativos a la disciplina Informática en la mayoría de los países occidentales. En contraposición, en algunos países orientales la participación de las mujeres en las TIC, muestra porcentajes elevados, ejemplo de ello son países como Taiwán, Filipinas y Mauricio (Fan T et al, 2004).

La progresiva y alarmante disminución de mujeres inscriptas en carreras STEM¹ a nivel global, han provocado la inquietud de diversos organismos internacionales que destacan la imperiosa necesidad de promover el estudio de estas carreras entre las jóvenes (Pavez I, 2015) (AAUW, 2015). Entre estos organismos, vale destacar el programa “STEM and Gender Advancement” de UNESCO (“GENDER AND SCIENCE”, s.f.) y la iniciativa “Girls in ICT” (Girls in ICT Portal, s.f.) de la ITU². Estas políticas orientadas a una mayor participación de las mujeres en las TIC, se alinean con las definidas por la UNLP. Evidencia de ello es que este año, por primera vez, el 100% de los/as estudiantes que ingresaron a la UNLP contaron con la posibilidad de acceder a un espacio introductorio de formación en género, coordinado por la Prosecretaría de Derechos Humanos de la UNLP. Esta iniciativa forma parte del eje promocional-preventivo que contiene el “Programa Institucional contra las Violencias de Género” de esta Universidad (Prevención de Violencias de Género, s.f). Verónica Cruz, prosecretaria de Derechos Humanos de la UNLP, recordó que “la finalidad es ofrecer una aproximación conceptual a la temática de género, mediante una estrategia participativa que propicie la problematización y desnaturalización de discursos y prácticas sexistas y machistas que obturan la construcción de relaciones igualitarias y respetuosas en el ámbito universitario”. En la Facultad de Informática, la intervención fue incorporada al Taller de Introducción a la Vida Universitaria (TIVU) que es parte del curso de Ingreso a las carreras, resultando en una oportunidad para concretar el primer contacto con las alumnas ingresantes, con la expectativa de establecer un vínculo que es imprescindible para elaborar diagnósticos según los distintos ejes

de la línea de investigación que se propone. Para mostrar en números, en el ingreso 2018 a las carreras de la Facultad de Informática de la UNLP, la distribución de mujeres y varones fue la que se detalla en el cuadro a continuación.

Carreras	Mujeres	Varones	%Mujeres
APU	53	230	18,73
ATIC	51	239	17,59
LI.	43	217	16,54
LS	49	229	17,63

Referencias: APU: Analista Programador Universitario; ATIC: Analista TIC; LI: Licenciatura en Informática y LS: Licenciatura en Sistemas

La escasa participación de mujeres en carreras tecnológicas advierte una inequidad que se traduce en barreras que impiden el desarrollo completo del potencial, generando disparidades que se reflejan en el acceso a oportunidades laborales en el sector TIC entre varones y mujeres, siendo aún mayor esta brecha en países en desarrollo (Pavez, I. 2015).

El LINTI viene realizando actividades y encuentros vinculados a la temática de “Género en las TIC” desde hace varios años. Ejemplo de ello es el evento “Mujeres y TIC” realizado en los años 2016 y 2017 en conmemoración del “Día Internacional de las Mujeres en TIC” (s.f.), acompañando la iniciativa “Girls in ICT” de la ITU. Este encuentro está organizado en diferentes estaciones-taller, de las que participan estudiantes y docentes de la Facultad de Informática sobre temáticas que muestran “qué se puede hacer con las tecnologías digitales” y charlas de mujeres referentes que sean motivadores para el público adolescente, como mujeres emprendedoras en el campo de videojuegos o comunidades de mujeres programadoras, entre otras. El evento está destinado a alumnas y alumnos de las escuelas secundarias de la región que concurren junto con sus docentes.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Proponer una línea de investigación acerca de la brecha de género en las carreras TIC, exige analizar la situación anterior y posterior a la vida universitaria, lo que se puede describir en preguntas tales como: ¿por qué las mujeres eligen/no eligen estudiar carreras TIC?, ¿qué ocurre cuando las han elegido y están en el ámbito universitario?, ¿cómo se da luego su inserción en la vida laboral y crecimiento profesional?. Cada una de estas etapas

¹ STEM: abreviatura de Science, Technology, Engineering and Mathematics

² ITU: International Telecommunications Union
26 y 27 de Abril de 2018

conlleva la definición de políticas, estrategias y mecanismos institucionales para su abordaje e indicadores propios para habilitar su evaluación. El abordaje de los temas de género, en los procesos de formación universitaria, se expresan en las propuestas educativas y de formación formal y no formal de los sujetos involucrados. En este sentido las especificidades de la disciplina Informática, como sucede en otras disciplinas, orientan los recortes de los conocimientos como también las problemáticas que se trabajan en las instituciones tanto a nivel del aula como en otros espacios de formación. Esta línea de investigación tensiona a los sujetos y a las instituciones interpelándolas en sus estructuras *instituidas* e *instituyentes*, es decir, en aquello que está, se encuentra establecido, no cuestionado y se reconoce como natural de la lógica institucional, y por otro lado lo que se está gestando, que según los aportes de Eduardo Remedi es lo *instituyente*, algo nuevo que cuestiona la identidad de los sujetos que han sido sostenidos históricamente. En esta línea de pensamiento se describen los ejes que definen las diferentes líneas de trabajo (Remedi E, 2004).

Eje de atracción a carreras TIC: este eje de indagación profundizará sobre la problemática de las niñas y adolescentes en las TIC. Este abordaje permite trabajar con anterioridad al ingreso a la Universidad. Conocer las condiciones y posibilidades de las jóvenes acerca del acceso a las TIC resultará sumamente útil para relacionar con la elección de una carrera universitaria del campo Informático. Este objetivo se sostiene en la definición de una estrategia metodológica que atraiga a las alumnas de nivel secundario y mujeres adolescentes que participen de espacios de formación no formal, y amplíe el universo de posibilidades para que las carreras del sector TIC se constituyan en una opción válida en su elección. La emergencia de la sociedad del conocimiento tiene como aspecto central la relevancia que adquiere el conocimiento experto y en este plano las carreras TIC tienen mucho para aportar. Las oportunidades laborales son algunos de los puntos que habitualmente se cita como atractivo, y que se constituye en la garantía de independencia económica de la mujer, siendo imprescindibles para salir de ciertos ámbitos de opresión y sometimiento.

Es decisiva la incorporación de estudiantes universitarias e investigadoras jóvenes para el contacto directo con la audiencia objetivo: ellas ponen en evidencia la viabilidad de la

propuesta.

Eje sobre procesos de formación en las carreras de la Facultad de Informática, ingreso, permanencia y graduación: este eje de indagación interpela a los dispositivos de formación que se despliegan, ya sea formales e informales, como también permitirá relevar e identificar vacancias relacionadas con esta problemática. En una mirada prospectiva permitirá pensar estrategias para la incorporar contenidos y acciones socialmente significativas que contribuyan con la formación de los/as jóvenes informáticos/as conforme con esta línea de trabajo.

Si bien el momento de inscripción e ingreso a las carreras de la Facultad representa un período corto, la comunicación y contención de las aspirantes en este período es decisiva. De acuerdo con las cifras mostradas en párrafos anteriores, la cantidad de mujeres es extremadamente baja con respecto a los varones, inclusive con un leve descenso respecto de años anteriores. Esa situación puede ser intimidante para las alumnas y de por sí, reforzando el estereotipo de carrera masculina y provocar el abandono y frustración. Asimismo merece considerarse la actitud del docente y tutor en esta etapa, determinadas situaciones, micromachismos³ incorporados culturalmente en el trato con las estudiantes, pueden atentar contra la permanencia de las mismas. De acuerdo a las experiencias que se ha tenido en el TIVU, uno de los grandes desafíos es trabajar para romper con los estereotipos que las y los jóvenes tienen incorporados, naturalizados, y que se manifiestan en este espacio. Este diagnóstico inicial sobre la realidad de las mujeres en las TIC nos convoca a trabajar de manera temprana y reconocer que este problema de investigación se proyectará y será un insumo valioso para trazar estrategias, acciones y generar dispositivos que contribuyan a incrementar la presencia y participación de las mujeres en las carreras de la Facultad de Informática.

La definición y puesta en marcha de diferentes instrumentos, el acompañamiento de los/as ingresantes y la detección de las potenciales razones de abandono o bajo rendimiento, es uno de los desafíos de esta etapa. La implementación de tutorías, actividades específicas con alumnas y docentes de la Facultad, un sistema de becas orientado a mujeres estudiantes y la

³ micromachismos: Es un término acuñado por el terapeuta argentino Luis Bonino en 1990. Según Bonino, se trata de comportamientos masculinos que buscan reforzar la superioridad sobre las mujeres. "Son pequeñas tiranías, terrorismo íntimo, violencia blanda", "suave" o de baja intensidad, tretas de dominación, machismo invisible o sexismo benevolente".

realización de encuestas, son algunos de los instrumentos a accionar.

Eje de retención y potenciación del talento durante la carrera: en este eje se trabajará con las alumnas que se encuentran cursando las carreras de la Facultad de Informática, con el objetivo de potenciar su talento y favorecer un buen desempeño durante el desarrollo de su vida universitaria. Por ejemplo, la maternidad es una de las vivencias que se citan habitualmente como causante de abandono parcial o definitivo de la carrera. Proponer facilidades y extensiones en el caso de las alumnas embarazadas o para madres y padres con hijos en períodos de lactancia hasta el año inclusive, puede resultar de una importancia decisiva para el alumno y la alumna que transitan la maternidad/paternidad. Prácticas como los programas de tutorías y el seguimiento del desempeño de las alumnas, y el contacto con mujeres profesionales en el ámbito de las TIC pueden resultar de una valorada ayuda en su vinculación con la enseñanza, la extensión, la investigación y el mercado laboral.

Eje de inserción laboral y crecimiento profesional: la inserción laboral de los y las alumnas de Informática habitualmente se inicia durante sus estudios, en etapas tempranas, debido a la gran oferta laboral existente en el sector TIC tanto desde el ámbito privado como público. La alumna que estudia y trabaja puede sumar experiencia y conocimiento, lo cual es un importante valor agregado en su formación profesional. Sin embargo, hay que tener en cuenta que si a su dedicación al trabajo se suma la maternidad, la continuidad de los estudios se ve amenazada. La toma de conciencia acerca de roles no estereotipados en la vida adulta (CEPAL, 2011), es un instrumento que promueve la defensa de sus derechos por parte de la mujer. La responsabilidad compartida en la crianza de los/as hijos/as, la demanda de espacios y tiempos para la lactancia, el percibir el mismo sueldo que un hombre ante igual tarea, son algunos de los derechos que deben ser exigidos y garantizados.

Indagar acerca de la realidad de las mujeres en el mundo laboral del sector TIC mediante diferentes instrumentos (charlas, entrevistas, prácticas rentadas, becas) nos aporta un conocimiento que podemos transmitir a las alumnas que transitan las primeras experiencias en el mundo del trabajo. En cuanto a la dedicación a la investigación, establecer semilleros de investigadoras en pregrado-postgrado, brindar oportunidades de becas de iniciación científica y estrategias

orientadas a apoyar publicaciones y/o difusión de resultados, podría resultar en acciones válidas para entusiasmar a las alumnas en este campo.

La línea de investigación “Género en las TIC” propone un abordaje de estos ejes mediante un enfoque interdisciplinario, con la interacción con investigadores del campo de las Ciencias de la Educación y Psicología, de nuestra Universidad. Se prevé la participación conjunta con el área de extensión y de docencia, como una contribución integrada al objeto de investigación.

3. RESULTADOS Y OBJETIVOS

Si bien el equipo de trabajo que conforma esta línea de investigación ha incursionado en actividades de investigación que nos han permitido identificar algunos indicadores cuantitativos (Díaz J, 2008) (Díaz J, 2013), se trata de una línea de investigación incipiente en el LINTI. En esta etapa iniciática se hace imprescindible la conformación de un equipo interdisciplinario para el abordaje integral de la problemática de género en las TIC, que permita explicar las causas de la escasa participación de las mujeres en carreras TIC de la Facultad de Informática de la UNLP. El desafío de esta línea de investigación es trabajar, profundizar y producir conocimientos sobre la problemática de género vinculada a la enseñanza y los procesos de formación de estudiantes de Informática en todas sus dimensiones. La preocupación se centra en detectar la emergencia de problemáticas propias y repensar la dimensión de la formación de los sujetos. Los objetivos de esta línea de trabajo son:

- Identificar los factores relacionados a la elección de las carreras TIC en adolescentes y jóvenes: cuáles son sus representaciones sobre el ejercicio profesional, sobre la disciplina Informática y su autopercepción de competencias.
- Elaborar un diagnóstico que permita explicar la escasa participación de las mujeres en carreras TIC en la Facultad de Informática de la UNLP.
- Promover relaciones igualitarias y respetuosas en la Facultad de Informática de la UNLP que favorezca la permanencia de las mujeres en las carreras y su inserción en el campo laboral.
- Interactuar con líneas de trabajo similares en las otras carreras STEM de la UNLP, para la realización conjunta de diagnósticos y la replicabilidad de experiencias, cuando sea posible.

- ~~Reconocer la importancia y dimensión de~~
RedUNCI - UNNE - ISBN 978-987-3619-27-4

la problemática para la elaboración de diferentes propuestas de concientización entendiendo que los/as estudiantes son una fuente de formación de recursos y formadora de pares.

Para llevar adelante estos objetivos se trabajará, en una primera etapa, con escuelas secundarias y espacios de educación no formal de la región de La Plata, Berisso y Ensenada, en el diseño y piloteo de experiencias de creación con medios digitales que permitan acercar las ideas del pensamiento computacional a las adolescentes (Wing J, 2006). A partir de estas intervenciones se recolectarán datos pre y post intervención. Por otro lado, se trabajará con estudiantes ingresantes y regulares de la Facultad de Informática en actividades que permitan identificar sus propias percepciones sobre la elección de las carreras para analizar los factores que atentan contra la participación de mujeres. Asimismo el equipo de trabajo ha comenzado a integrarse a iniciativas similares en el ámbito regional, nacional e internacional, ejemplo de ello es la participación en Benchmarking organizado por Universidad Columbus (“U-Benchmarking Club: Mejorando la participación de Mujeres en programas STEM”, 2017).

Los resultados esperados en esta etapa se relacionan con contar con un diagnóstico en cuanto a la elección de carreras TIC por parte de las mujeres, la retención y el potenciamiento del talento de las alumnas y la situación en el ámbito laboral, que conducirán a definir y concretar acciones cuyos resultados pueden ser evaluados en su efectividad según indicadores predefinidos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo que lleva adelante estas líneas son docentes investigadores/as del del LINTI. Asimismo, se están comenzando a formular tesinas de grado, tesis de postgrado, proyectos de extensión y actividades de cátedras relacionadas con a la problemática presentada.

5. REFERENCIAS

AAUW (2015). *Solving the Equation: The Variables for Women's Success in Engineering and Computing*. American Association of University Women Educational Foundation.

Castaño C. (2008). *La segunda brecha digital*. Madrid: Ediciones Cátedra.

- (2009). *La segunda brecha digital y las mujeres jóvenes*. Quaderns de la Mediterrània, 11, 218-224.

CEPAL (2011). TIC y género. Newsletter eLAC nº 16. Septiembre.

Díaz J., Osorio M., Amadeo A. (2008). *Trends in studying IT degrees at National University of La Plata – Argentina. A gender analysis*. Women in Engineering & Computer Science.

Díaz J., Queiruga C., Molinari L., Amadeo A., Osorio M. (2013). *Mujeres y Tecnología: evaluación de la participación en las actividades de la Facultad de Informática de la UNLP*. VI Congreso de la Mujer Latinoamericana en la Computación.

Día Internacional de las Mujeres en TIC (s.f.). Disponible: <http://bit.ly/2G3LioN>; <http://bit.ly/2DCZ0K6>

Girls in ICT Portal (s.f.). Disponible: <http://bit.ly/2pkzmUP>

Gürer, Denisse y Camp, Tracy (2002). *An ACM-W literature review on women in computing*. SIGSCE Bulletin, 34 (2), 121-127.

Fan Tai-Sheng, Li Yi-Ching (2004). *Gender issues and computers: college computer science education in Taiwan*. Computers & Education. 44, 285-300.

Pavez, Isabel (2015). *Niñas y mujeres de latinoamérica en el mapa tecnológico: una mirada de género en el marco de políticas públicas de inclusión digital*. Cuaderno SITEAL, UNESCO, IPEE-UNESCO, OEI.

Prevención de Violencias de Género (s.f) Disponible: <http://bit.ly/2pmfPVj>

Remedi, E. (2004). *La institución: Un entrecruzamiento de textos*. En Remedi, E. (coord.) Instituciones Educativas. Sujetos, historias e identidades, pp. 25-55. México: Plaza y Valdés Editores.

U-Benchmarking Club: Mejorando la participación de Mujeres en programas STEM (11 de octubre de 2017). Disponible: <http://bit.ly/2xyUWfy>

UNESCO (2013). *Situación Educativa de América Latina y el Caribe: Hacia la educación de calidad para todos al 2015*. En Cristián Bellei (coordinador) del Centro de Investigación Avanzada en Educación de la Universidad de Chile.

UNESCO (2005). *Feminización de la matrícula de educación superior en América Latina y el Caribe*. Rosaura Sierra; Gisela Rodríguez, comp. México D.F.: UNESCO IESALC.

Wing, J. (2006). *Computational Thinking*. Communications of the ACM, vol. 49, 33-35.

ANALIZANDO EL IMPACTO DE TALLERES DE PROGRAMACIÓN EN ESCUELAS CON RESPECTO AL INGRESO DE ALUMNOS EN CARRERAS DE INFORMÁTICA

Ana Garis, M. Claudia Albornoz, Mario Silvestri
Universidad Nacional de San Luis, San Luis, Argentina
{agaris,albornoz,msilvestri}@unsl.edu.ar

RESUMEN

La falta de vocaciones en carreras universitarias relacionadas a la Informática es una problemática a nivel mundial. Para enfrentar este problema se han aplicado diferentes estrategias, tales como el dictado de talleres sobre temáticas afines, en establecimientos educativos, especialmente a nivel secundario. La presente línea de investigación se enfoca en la evaluación del impacto que han tenido dichos talleres con respecto al objetivo de incrementar la matrícula de alumnos ingresantes a carreras vinculadas con la Informática, en particular, en la Universidad Nacional de San Luis.

Palabras Claves

Enseñanza de la Programación, Escuelas secundarias, Difusión de Carreras de Informática

CONTEXTO

Este trabajo se encuentra enmarcado dentro del Proyecto P-031516, cuyo título es: "Ingeniería de Software: conceptos, prácticas y herramientas para el desarrollo de software de calidad", de la Facultad de Ciencias Físico, Matemáticas y Naturales, de la Universidad Nacional de San Luis. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se han logrado importantes vínculos con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el Programa de Incentivos.

1. INTRODUCCIÓN

El escaso interés de los jóvenes en el área Informática y las Ciencias de la Computación es una problemática a nivel internacional. Diversas organizaciones tanto públicas como privadas han presentado propuestas para afrontar este problema. Una de las modalidades utilizadas ha sido el dictado de talleres en escuelas sobre áreas asociadas a la computación, tales como la programación. En este contexto desde la Universidad Nacional de San Luis (UNSL), desde 2014 a 2016 inclusive, se dictaron talleres de Programación a alumnos de los últimos años de escuelas secundarias de la ciudad de San Luis [1], en el marco del programa institucional denominado "Articulación Universidad y Escuela Secundaria". Dicho programa tenía como fin divulgar contenido científico entre alumnos de escuelas secundarias, para despertar vocaciones tempranas en ciencias exactas y naturales. El objetivo primario de los talleres de Programación fue brindar a los estudiantes secundarios de los últimos años un acercamiento a la programación y fomentar el estudio de carreras de Informática. En el ámbito de la UNSL, esto representa la elección de alguna carrera correspondiente al Departamento de Informática [2]. Los talleres siguieron como guía las propuestas presentadas en "La Hora del Código" [3], una iniciativa que promueve el aprendizaje de la programación, incluyendo jóvenes en edad escolar.



Fig. 1: Actividades de La Hora del Código

La Hora del Código es impulsada por Code.org, una organización que cuenta con una completa plataforma online. La plataforma presenta múltiples actividades de diferente nivel de complejidad. La Fig. 1 muestra algunas de las actividades disponibles actualmente. Las propuestas tienen como protagonistas a personajes populares de la comunidad juvenil, y emplean como estrategia didáctica al juego, esto atrae la atención de los estudiantes y les permite mantener la concentración durante el desarrollo de los talleres [4,5,6]. Estudios recientes han mostrado que la utilización del material de Code.org en las escuelas, serviría para despertar vocaciones en Informática, en particular alumnos de sexo femenino que son quienes muestran menor interés en esta área [7].

Los talleres incluían una charla motivacional final, en donde se mostraba la oferta académica que ofrece la UNSL, y la entrega de cuestionarios para consultar el interés por el área de Informática y las Ciencias de la Computación. Esta información resultó de gran importancia para llevar a cabo un análisis estadístico, observando que los estudiantes de cursos inferiores, tales como 4to. año, expresaron

mayor interés que los de los últimos años [1]. Sin embargo, la elección final de estos alumnos, plasmada en la inscripción formal posterior en la Universidad, no ha sido estudiada aún.

La presente línea de investigación propone estudiar el ingreso a carreras de Informática de la UNSL de alumnos participantes en los talleres de Programación. El trabajo está dirigido a evaluar el impacto real que tuvieron los talleres en cuanto al objetivo principal que éstos tenían: despertar vocaciones para promover el estudio universitario de carreras vinculadas a la Informática y Ciencias de la Computación.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los ejes de investigación se basan en los siguientes puntos.

- Análisis de antecedentes que describan cómo se ha evaluado el ingreso de alumnos a la universidad, considerando si éstos participaron previamente en talleres brindados por la institución.
- Especificación del proceso que debe transitar un alumno, desde el momento que se inscribe en la carrera hasta finalizar el 1er. año.
- Recolección de datos informales registrados por los docentes de cátedras del primer año de carreras del Departamento de Informática.
- Recolección de datos estadísticos brindados por los sistemas informáticos de la UNSL.
- Análisis de datos y publicación de resultados obtenidos. Se pretende hacer llegar los resultados a los establecimientos educativos en

donde se dictaron los talleres, como una forma de reforzar los vínculos entre escuela y universidad.

3. RESULTADOS OBTENIDOS/ESPERADOS

Luego de llevar a cabo un análisis preliminar de antecedentes, se ha observado que existe escasa información sobre la evaluación del ingreso de alumnos a la universidad, en función de su participación previa en talleres brindados por la institución. Si bien, es posible encontrar en la bibliografía trabajos que exponen experiencias en el dictado de talleres de Programación en las escuelas [8,9,10], los mismos están más orientados a mostrar las estrategias didácticas asociadas a la enseñanza de la Programación y no al impacto en cuanto al ingreso posterior de los alumnos en universidades.

Se ha realizado la especificación informal del proceso que debe transitar un alumno, desde el momento que se inscribe en la carrera hasta finalizar el 1er. año. Éste incluye un curso de ingreso sobre matemáticas. Dependiendo la aprobación del mismo, los alumnos pasan al trayecto normal o van a otro trayecto de formación con apoyo. Del listado de ingresantes que rinden el curso de matemáticas se ha podido observar que reprueba más del 50%. Se dispone actualmente del listado de alumnos y sus escuelas de origen brindados por docentes de algunas cátedras del primer año de carreras del Departamento de Informática, durante el periodo 2015-2017. Estos datos deberían ser contrastados con el registro de estudiantes participantes en los talleres.

Como trabajo futuro se debe recoger información sobre el resto de los cursos de 1er. año. Para esta tarea se requerirá acceder a los datos disponibles en los sistemas informáticos de la UNSL. Queda pendiente también el análisis y publicación de los resultados. Los mismos permitirán no solo sacar conclusiones sobre la eficacia

y eficiencia de los talleres en relación al objetivo de incrementar el número de alumnos ingresantes, sino también detectar problemáticas inherentes del transitar del estudiante de 1er. año de las carreras de Informática.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro del proyecto, donde se encuentra inserta ésta línea de investigación, se están llevando a cabo diferentes tesis de grado y posgrado. Éstas últimas correspondientes a la Especialización en Ingeniería de Software, Maestría en Ingeniería de Software y al Doctorado en Ingeniería Informática.

Referencias

- [1] Garis A., Albornoz C., Silvestri M.. “La Hora del Código: Promoviendo la Programación en Escuelas Secundarias de San Luis”, Proceedings of the 5to. Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI 2017), pp. 1215-120, 2017.
- [2] Departamento de Informática (UNSL), sitio web <http://www.dirinfo.unsl.edu.ar/>
- [3] Code.org, sitio web <http://code.org/>
- [4] Charsky D., “From Edutainment to Serious Games: A Change in the Use of Game Characteristics”, Journal Games and Culture, Vol. 5, No. 2, pp. 177–198, 2010.
- [5] Frittelli V., Tartabini M., Teicher R., Steffolani F., Serrano D., Fernández J., Bett G., Strub A. “Desarrollo de Juegos como Estrategia Didáctica en la Enseñanza de la Programación”, in Proceedings 3er Congreso Nacional de Ingeniería

Informática /Sistemas de Información (CoNaIISI 2015), Buenos Aires, Argentina, 2015.

[6] Sanz, C. V. “Aprender a programar en tiempos digitales”, *Bit & Byte*, Año 2, No. 4, pp. 12-14, 2016.

[7] Kalelioğlu F., “A new way of teaching programming skills to K-12 students: Code.org”, *Journal Computers in Human Behavior*, Elsevier, Vol. 52, 2015, pp. 200-210.

[8] Queiruga, C., Banchoff C., Martín S., Rosales V. A., López F., “PROGRAMAR en la Escuela”, in *Proceedings XVIII Workshop de Investigadores en Ciencias de la Computación (WICC 2016)*, Entre Ríos, Argentina, 2016.

[9] Program.AR, sitio web <http://program.ar>

[10] Díaz, J., Banchoff, C., Queiruga, C., Martín, E., “Experiencias de la facultad de informática en la enseñanza de programación en escuelas con software libre”, in *Proceedings Congreso Iberoamericano De Ciencia, Tecnología, Innovación y Educación*. Buenos Aires, Argentina, 2014

ESTRATEGIAS INNOVADORAS DE ENSEÑANZA DE LA PROGRAMACIÓN Y DIDÁCTICAS ESPECÍFICAS PARA FOMENTAR EL PENSAMIENTO COMPUTACIONAL

Gladys N. Dapozo, Cristina L. Greiner, Raquel H. Petris, María C. Espíndola, Ana M. Company, María F. Piragine, Marta Stopello, Irene Lucero

Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura
Universidad Nacional del Nordeste. 9 de Julio N° 1449. Corrientes. Argentina.

{gndapozo, cgreiner, rpetris}@exa.unne.edu.ar, mcespindola@yahoo.com,
anamacom@hotmail.com, mafepiragine@hotmail.com, mstopello@hotmail.com,
irmairenprof@gmail.com

RESUMEN

La formación en ciencia, tecnología, ingeniería y matemáticas se considera clave para resolver muchos de los desafíos actuales y futuros de la humanidad, por lo cual la tendencia mundial en educación es fomentar el pensamiento computacional y la programación, especialmente en los niveles educativos preuniversitarios. En este proyecto se plantean tres líneas principales de investigación, una enfocada en estrategias educativas que incorporen métodos y herramientas innovadoras para la enseñanza de la programación en las carreras de Informática, otra orientada a las didácticas específicas para actualizar la formación de los profesores del campo de las Ciencias de la Computación, y finalmente, una línea enfocada en definir estrategias para salvar los obstáculos epistemológicos surgidos en el proceso de enseñanza y aprendizaje de las asignaturas básicas (Física, Química, Matemática) en el primer año de las carreras de Ciencias Exactas, vinculados en forma directa con los preocupantes índices de desgranamiento y de abandono de las carreras STEM (Science, Technology, Engineering and Mathematics).

Palabras clave: Enseñanza de la programación. Didácticas específicas. Pensamiento computacional. Carreras STEM.

CONTEXTO

Las líneas de I/D corresponden al proyecto 16F018 “Promoción del pensamiento computacional para favorecer la formación en STEM”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE), iniciado en el año 2017.

1. INTRODUCCIÓN

Los inconvenientes a los que se enfrentan las carreras de formación en disciplinas que involucran saberes en STEM (Science, Technology, Engineering and Mathematics), están siendo estudiados y analizados en diversas universidades que se ven afectadas por una marcada disminución de estudiantes en este tipo de carreras. Algunas de las problemáticas son: la constante disminución en la matrícula de estas disciplinas, a su vez el marcado desgranamiento en los primeros años y el reducido porcentaje de matrícula femenina. Un informe del Consejo Presidencial de Asesores en Ciencia y Tecnología (PCAST) en EE.UU. señala que se requerían, en la próxima década, aproximadamente 1 millón de graduados universitarios en los campos de STEM. Además, en ese país menos del 40% de los estudiantes que ingresan a la universidad con la intención de especializarse en un campo STEM lo logran [1].

El estudio de Katz [2] señala que se aprecia una producción limitada de

graduados en STEM en relación al número de egresados respecto de las necesidades del campo productivo. En particular destaca la falta de masa crítica en capacidades para trabajar en el área de la innovación digital.

Esta formación en ciencia, tecnología, ingeniería y matemáticas se considera clave para resolver muchos de los desafíos actuales y futuros de la humanidad, por lo cual la tendencia mundial en educación es fomentar el pensamiento computacional y la programación, especialmente en los niveles educativos preuniversitarios. Con lo cual un problema que se plantea es la actualización de conocimientos, estrategias didácticas y herramientas de enseñanza de programación para los profesores de asignaturas vinculadas con las ciencias de la computación. Muchos docentes no tienen las habilidades profesionales en informática o pensamiento computacional, o no tienen presente nuevos enfoques de la didáctica de la programación [3].

Ante este problema de carácter global, existen numerosos programas que proponen soluciones, que buscan mejorar la educación en estos temas desde los primeros niveles de educación (“Some STEM for All”), concentrar los esfuerzos sólo en los interesados o en los destacados (“All STEM for Some”) [4], fortalecer la formación de los docentes de STEM y promover el incremento de los mismos mediante beneficios adicionales [5].

Por otra parte, se sostiene que todos los niños deben tener la posibilidad de acceder a la enseñanza de computación en su vida escolar. Las habilidades que desarrollan los alumnos al ser educados en ciencias de la computación exceden lo referido estrictamente a este tema. Estas habilidades, vinculadas con el “Computational Thinking” propuesto por Wing [6], cumplen un rol de creciente importancia en la educación moderna.

Por tanto, cuando se trata de alfabetización digital, esta debe integrar también nociones sobre los lenguajes de las computadoras. En este marco, tanto la programación como el pensamiento computacional resultan relevantes para el aprendizaje. Al comprender su semántica y su lógica en la resolución de problemas, los alumnos también se preparan para entender y cambiar la realidad [7].

En las carreras de Informática la problemática del aprendizaje de la programación continúa siendo una preocupación vigente. En los planes de estudio, generalmente, estos contenidos están en las primeras asignaturas, con lo cual, a las dificultades propias de programación, se agregan las características de los estudiantes que recién se inician en la vida universitaria. Sin embargo, los estudios realizados señalan que la problemática no reside en la incapacidad para resolver problemas propiamente dicha sino en el escaso desarrollo del pensamiento computacional y de los procesos de abstracción que ellos requieren [8].

Por tanto, de la literatura se desprende que fomentar el pensamiento computacional favorece a los alumnos que realizan una formación específica en Ciencias de la Computación, como así también, a niños y jóvenes que no necesariamente realizarán una formación profesional en el campo de la Informática.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación de este proyecto, están enfocadas en:

- a) Estrategias educativas que incorporen métodos y herramientas innovadoras para la enseñanza de la programación en las carreras de Informática.
- b) Didácticas específicas para actualizar la formación de los profesores del campo de las Ciencias de la Computación. Esta línea busca

acompañar las políticas públicas, en particular la iniciativa Program.Ar, para lograr promover “cambios de fondo en la enseñanza en escuelas primarias y secundarias de varios temas relacionados con la computación, convencidos de que son un elemento clave para que el país pueda aprovechar las enormes oportunidades que brindan estas tecnologías”. Coincidiendo que aprender la “verdadera computación” (las ciencias de la computación) será muy beneficioso para que todos los alumnos argentinos desarrollen habilidades y competencias fundamentales para la vida moderna [9].

- c) Estrategias para salvar los obstáculos epistemológicos surgidos en el proceso de enseñanza y aprendizaje de las asignaturas básicas (Física, Química, Matemática) en el primer año de las carreras de Ciencias Exactas, vinculados en forma directa con los preocupantes índices de desgranamiento y de abandono de las carreras STEM (Science, Technology, Engineering and Mathematics),

3. RESULTADOS OBTENIDOS/ESPERADOS

En las carreras universitarias de Informática generalmente la enseñanza de la programación aparece al inicio del proceso formativo, con lo cual, a las dificultades de adaptación de los estudiantes a las exigencias de la vida universitaria se añaden las dificultades propias del aprendizaje de la programación. En [10] se describe una modalidad de introducción a la programación implementada en la asignatura Algoritmos y Estructura de Datos I de la Licenciatura en Sistemas de Información (LSI), que consiste en el desarrollo de un conjunto de actividades, basadas en herramientas lúdicas, con el objetivo de estimular el pensamiento computacional e incrementar la

motivación de los alumnos. Para el desarrollo de las actividades se propuso un método de resolución de problema que incorpora herramientas conceptuales, como la abstracción y la modularización, y herramientas del lenguaje, como las estructuras de control repetitivas, alternativas y el uso de parámetros. En particular, se insistió en los conceptos de “abstracción”, “descomposición del problema en partes”, y en la “legibilidad” de la solución. Los resultados indican que estas actividades contribuyeron positivamente al incremento de la motivación de los estudiantes y a la incorporación de un método de resolución de problemas que facilita la transición a la programación con lenguajes de programación convencionales.

Además, incorporar estrategias de enseñanza de programación para alumnos universitarios de carreras de Informática utilizando herramientas lúdicas incrementa la motivación de los docentes y alumnos dado que el aprendizaje de los conceptos básicos se da en un contexto ameno y recreativo. Consolidar de esta forma un método de resolución de problema que se mantenga aun cambiando la herramienta ofrece un marco de referencia que facilitará la programación con herramientas convencionales [11].

En este contexto de dificultades reconocidas sobre la enseñanza de la programación, se realizó un estudio acerca de características comunes que tienen los alumnos que logran sortear las dificultades y aprender a programar. Conocer cómo actúan los alumnos que se han destacado por su buen desempeño aporta información para mejorar las estrategias de enseñanza y reducir el número de alumnos que queda libre por parciales, o peor aún, que abandonan el cursado. En [12] se muestran los resultados de este estudio.

En la línea de capacitación en programación, los resultados de la experiencia de formación en Didáctica de la Programación destinada a docentes de los niveles no universitarios de la ciudad de Corrientes se publicaron en [13].

En cuanto a Resultados Esperados, el proyecto continuará con las siguientes actividades:

a) Análisis del impacto de las acciones de promoción de la programación en las escuelas.

Dado que, en los últimos 3 años, por diferentes medios, los alumnos de las escuelas secundarias tuvieron la oportunidad de acceder a conceptos de programación con herramientas visuales, interesa conocer si esta experiencia previa redundó positivamente en el aprendizaje de la programación en el inicio de una carrera de Informática. Para este estudio se considerarán los alumnos ingresantes a la LSI de la UNNE, 300 aproximadamente.

b) Análisis de la motivación de los docentes para incorporar la programación en el aula, en función del perfil del docente y del nivel educativo en el cual se desempeña.

En el marco de la actualización de los docentes, mediante el dictado del curso Programación y su Didáctica, que promueva la Fundación Sadosky se evaluarán intereses, expectativas y posibilidades reales de incorporación de contenidos de Ciencias de la Computación de los docentes. Por otra parte, indagar también acerca de las propuestas/métodos/herramientas que los docentes consideran más apropiadas para el nivel educativo en el que se desempeñan,

c) Evaluación de las estrategias didácticas para la enseñanza de las asignaturas básicas (Física, Química, Matemática) en asignaturas de primer año de las carreras de Ciencias Exactas y elaboración de propuestas didácticas superadoras.

4. FORMACIÓN DE RECURSOS HUMANOS

En el Grupo de Investigación GI-TIPC (Tecnologías Informáticas y Pensamiento Computacional) están involucrados ocho docentes investigadores que provienen de distintas áreas de la Facultad de Ciencias Exactas de la UNNE, Informática, Matemática y Física, con el objetivo de darle la mirada interdisciplinaria que esta problemática requiere. En el marco del proyecto dos tesis de posgrado desarrollan su trabajo final de la Maestría en Tecnologías de la Información de la UNNE.

5. REFERENCIAS

- [1] President's Council of Advisors on Science and Technology. (2012). "Report to the president. Engage to excel: producing one million additional college graduates with degrees in science, technology, engineering, and mathematics". Disponible en: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf
- [2] Katz, R. (2016). TIC, digitalización y políticas públicas. En Entornos Digitales y Políticas Educativas. IIPE-UNESCO.
- [3] Min Xiao, Xiaohua Yu. A Model of Cultivating Computational Thinking Based on Visual Programming. The Sixth International Conference of Educational Innovation through Technology. 2017.
- [4] Atkinson, R.D.; Mayo, M. "Refueling the U.S. innovation economy: Fresh Approaches to Science, Technology, Engineering and Mathematics (STEM) Education". Disponible en: <http://www.itif.org/files/2010-refueling-innovation-economy.pdf>
- [5] Committee Members. "Rising above the gathering storm, revisited". Disponible en: <http://www.sandia.gov/NINE/documents/RisingAbove.pdf>

- [6] Jeannette M. Wing, Computational Thinking. Communications of the ACM Viewpoint, March 2006, pp. 33-35.
- [7] María Florencia Ripani. Competencias de Educación Digital. Ministerio de Educación de la Nación. Argentina. 2017. Libro digital. Disponible en: <http://planied.educ.ar/wp-content/uploads/2017/09/Competencias-05.pdf>
- [8] Valverde Berrocoso, J.; Fernández Sánchez, M.R; Garrido Arroyo, M. del C. El pensamiento computacional y las nuevas ecologías del aprendizaje. RED-Revista de Educación a Distancia, 46(3). 2015. Disponible en: http://www.um.es/ead/red/46/valverde_et_al.pdf
- [9] Sadosky. CC-2016 Una propuesta para refundar la enseñanza de la computación en las escuelas argentinas. 2013.
- [10] G. N. Dapozo, C. L. Greiner, R. H. Petris, M. C. Espíndola and A. M. Company, "Introduction to programming based on playful activities in the university," 2017 XLIII Latin American Computer Conference (CLEI), Cordoba, 2017, pp. 1-8. doi: 10.1109/CLEI.2017.8226437.
- [11] Gladys N. Dapozo, Cristina L. Greiner, Raquel H. Petris "Enseñar programación desde los juegos a lo formal". 1er Congreso de Educación y Tecnologías del Mercosur: de la digitalización a la virtualización. ISBN 978-987-3619-26-7. 2017.
- [12] Gladys N. Dapozo, Cristina L. Greiner, Raquel H. Petris, María C. Espíndola, Ana M. Company-Enseñanza de la Programación en la Universidad. Factores que Inciden en el Buen Desempeño de los Estudiantes. Anales CONAISI 2017. 5to Congreso Nacional de Ingeniería Informática / Sistemas de Información 2-3 Noviembre, Santa Fe - Argentina
- [13] Dapozo, G.; Petris, R.; Greiner, C. (2016). "Programación en las escuelas. Experiencia de formación docente en el Nordeste Argentino". Revista REMEIED: Memorias del Encuentro Internacional de Educación a Distancia. Desafíos de la cultura digital en América Latina. ISSN: 2395-8901. Diciembre 2016-noviembre 2017

Escuelas TIC: las tecnologías digitales en las aulas

Claudia Queiruga, Claudia Banchoff Tzancoff, Paula Venosa, Sofía Martín, Vanessa Aybar Rosales, Soledad Gomez, Isabel Kimura

Laboratorio de Investigación de Nuevas Tecnologías Informáticas (LINTI). Facultad de Informática. Universidad Nacional de La Plata

50 y 120. La Plata

{claudiaq, cbanchoff, pvenosa, vaybar, sgomez}@info.unlp.edu.ar, {smartin, ikimura}@linti.unlp.edu.ar

RESUMEN

En una sociedad cada vez más conectada, en la cual los objetos tecnológicos forman parte de nuestro cotidiano, educar en tecnología se torna imprescindible en la formación de niños/as y jóvenes. Si bien hay condiciones que no permiten asegurar que esto sea posible para todos los ciudadanos sin importar edad, localización y situación social, es necesario contemplar que la formación de nuestros niño/as y jóvenes en estas temáticas es una responsabilidad que afecta a gobiernos, instituciones educativas y es tema de debate en la sociedad en general.

La Informática y las TIC están insertas en la vida cotidiana y, aunque las currículas escolares tienden a incorporar el uso de TIC en las distintas áreas, hoy en día sigue observándose un uso instrumental de las mismas. “EscuelasTIC” es una línea de investigación que da continuidad a línea de trabajo “PROGRAMAR en la Escuela”, sumando a dicho trabajo aportes que van más allá de la enseñanza de la programación, y que abarcan al campo disciplinar de la Informática en la escuela.

Palabras clave: Informática, enseñanza de programación, robótica educativa, pensamiento computacional, formación docente, TIC.

CONTEXTO

La línea de investigación “EscuelasTIC” presentada en este trabajo, está inserta en el proyecto de investigación “Internet del Futuro: Ciudades Digitales Inclusivas, Innovadoras y

Sustentables, IoT, Ciberseguridad, Espacios de Aprendizaje del Futuro” del Programa Nacional de Incentivos a docentes-investigadores, que se desarrolla en el LINTI de la Facultad de Informática de la Universidad Nacional de La Plata (UNLP). Este proyecto está acreditado por la UNLP y financiado por partidas del presupuesto nacional. Asimismo, esta línea de trabajo articula los procesos de investigación con proyectos de extensión de la UNLP, de los que participan docentes-investigadores del LINTI, dándole el sustento de pertinencia local y regional a los resultados obtenidos. En estos proyectos se llevan a cabo una serie de actividades relacionadas con la capacitación, acompañamiento y creación de contenido para el ámbito educativo y están orientados al uso y creación con tecnologías digitales en el aula y, en especial a la incorporación de conceptos de programación en un sentido transversal.

La línea de investigación aquí presentada da continuidad a “PROGRAMAR en la Escuela: nuevos desafíos en las aulas” presentada en eventos anteriores (Queiruga C, 2017).

1. INTRODUCCIÓN

Nuestro cotidiano está rodeado de objetos tecnológicos, empezando por el celular que nos acompaña todo el tiempo, hasta semáforos y casas inteligentes, pasando por drones que nos sobrevuelan, ropa tecnológica (del inglés *wearable technology*), entre otros elementos que cada vez son más evidentes en nuestra vida diaria. Estos elementos nos modifican, aumentan nuestras capacidades cognitivas y es por ello que es necesario entender qué pasa allí adentro. En el campo educativo se plantea la formación de ciudadanos/as que puedan

comprender los lenguajes digitales, ubicándolos como sujetos críticos y creadores de innovaciones con tecnologías digitales, por sobre la pasividad y el mero consumo tecnológico.

El uso de TIC se ha ido incorporando a las prácticas educativas en los distintos niveles de la escolaridad obligatoria, inclusive en algunos de ellos la disciplina Informática forma parte de algunos diseños curriculares escolares. El nuevo diseño curricular de la educación primaria de la provincia de Buenos Aires (DGCyE/Diseño curricular para la educación primaria de la provincia de Buenos Aires, 2018), es un ejemplo de ello, el cual contiene un módulo sobre la inclusión de TIC que intenta incorporar en forma transversal el uso de tecnologías digitales en las distintas áreas curriculares. Asimismo existe un espacio curricular en la escuela secundaria donde algunos conceptos relacionados a las Ciencias Informáticas se incluyen en la materia NTIC (DGCyE/Diseño curricular para la educación secundaria ciclo superior, 2010).

La Informática como disciplina aún no ha sido legitimada en el ámbito de la educación escolar, aunque hay algunas experiencias donde esto comienza a transformarse, aún no se ha llegado a los consensos necesarios sobre la inclusión del campo en las currículas escolares a nivel federal. Las escuelas de educación secundaria técnica son las únicas que cuentan con trayectos formativos en Informática, de esta manera el estudio sobre la disciplina no se da en forma transversal, ni forma parte integral de los contenidos a los que acceden la mayoría de los estudiantes. Incorporar en los diseños escolares la disciplina Informática constituye un logro y un espacio de seguridad para el desarrollo de procesos cognitivos vinculados al razonamiento lógico que permite predecir, analizar y explicar, a la formulación de algoritmos, a la descomposición de problemas en partes más simples, a la abstracción para manejar la complejidad, a la generalización mediante el descubrimiento de patrones y similitudes y, a la evaluación. Estas ideas, que identifican al “pensamiento computacional” (Wing J, 2006) (Wing J, 2008), son ampliamente aplicadas en la resolución de problemas usando computadoras y en la comprensión de los sistemas más allá del espacio escolar.

La enseñanza de Informática en los sistemas educativos es actualmente una preocupación global. Países como Nueva Zelanda, Estonia, Japón, Finlandia y Reino Unido, entre otros, han actualizado sus currículas escolares incluyendo la enseñanza de la programación en las escuelas (National curriculum in England, 2013) (The New Zealand Curriculum on-line, s.f). Varios estados de los Estados Unidos también han implementado políticas activas en respuesta al respaldo de la industria tecnológica mediante el movimiento “LearnToCode”, liderado fundamentalmente por las iniciativas globales code.org y codecademy.

En los últimos años en nuestro país, un conjunto de iniciativas y políticas federales, han contribuido a consolidar una perspectiva que fortalece las razones por las cuales resulta necesario situar la enseñanza de la Informática en la escuela. Actualmente forma parte de la agenda de trabajo del Ministerio de Educación de la Nación y de varios ministerios provinciales. Ejemplo de ello es el surgimiento de programas de políticas educativas, tales como el proyecto “Program.AR” (“Program.AR”, s.f), la creación del “Programa Conectar Igualdad” y el “Plan Nacional Integral de Educación Digital” (PLANIED, sf), cuyo objetivo es “integrar la comunidad educativa en la cultura digital, favoreciendo la innovación pedagógica, la calidad educativa y la inclusión socioeducativa. Su misión consiste en que todos los estudiantes de la Argentina adquieran las habilidades necesarias para desenvolverse en el mundo actual y en la sociedad del futuro”, que planifica la incorporación de drones, robótica, minicomputadoras educativas y nuevos contenidos en 3000 escuelas argentinas.

La incorporación de estos recursos tecnológicos generan desafíos para los docentes y autoridades de las escuelas, dado que no se encuadra en planes de capacitación docente en herramientas digitales, sino que el desafío es trabajar en la creación de contenidos digitales en las distintas áreas curriculares. Sin embargo, es común encontrar que dentro de las propuestas didácticas, el uso de las computadoras y del equipamiento informático es instrumental. Desde este enfoque los/as alumnos/as se configuran simplemente como usuarios de aplicaciones específicas y servicios en un sentido irreflexivo. Esta situación nos

resulta problemática en tanto no se promueve un uso crítico sobre los medios digitales con los que interactúan los/as niños/as y jóvenes. En esta línea de pensamiento se ubica la creación con medios digitales y la formación de ciudadanos digitales responsables, que conozcan el impacto de sus acciones y de la información que comparten públicamente.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Esta línea de investigación está organizada en tres ejes: formación docente, desarrollo de materiales didácticos y diseño y puesta en acción de experiencias de enseñanza en escuelas primarias y secundarias. El equipo de trabajo es interdisciplinario está integrado por docentes-investigadores informáticos/as y especialistas en Educación del LINTI. El enfoque de las intervenciones está orientado a la adopción del *pensamiento computacional* en la escuela a través de la programación y otras áreas de la disciplina Informática y, a la construcción de contenidos con tecnologías digitales pertinentes al espacio escolar.

Se está trabajando con docentes y estudiantes de 12 escuelas secundarias y 6 escuelas primarias del distrito escolar de La Plata, Berisso y Ensenada. Las actividades que se realizan con los docentes, constituyen instancias de formación docente claves para el desarrollo de los contenidos del área en cada una de las materias que los mismos dictan. Se planifican encuentros acordados en los que se trabajan temas y materiales que son insumos de trabajo en sus clases. Se trata de espacios de formación específicos para docentes de los distintos niveles sobre las nuevas didácticas en relación a la Informática y las TIC (Díaz J, 2014).

En relación a la enseñanza de la programación en las escuelas, el uso de lenguajes de programación visuales basados en bloques e icónicos, concebidos en clave de los intereses de los niños/as y adolescentes, y la manipulación de objetos físicos, facilitan la incorporación de conceptos, prácticas y perspectivas propias del *pensamiento computacional*.

Las interacciones entre el mundo físico y el virtual, y los procesos de automatización que se

observan, se pueden introducir mediante el aprendizaje de la robótica e “Internet de las cosas” (IoT, por sus siglas en inglés), sustentado en la programación. La posibilidad de visualizar los efectos físicos producidos por los comandos programados, la creación de artefactos informáticos que colaboren con el bienestar general de nuestra sociedad, constituyen un recurso pedagógico sumamente potente y motivador para los niños/as y adolescentes. Explorar el campo de IoT habilita la introducción de conceptos sobre el funcionamiento de las redes, resultando sumamente útil en actividades de formación del pensamiento computacional. El uso seguro y responsable de las tecnologías digitales forma parte de las actividades en esta línea de investigación y sus distintos ejes, íntimamente vinculados a la formación en ciudadanía digital.

3. RESULTADOS Y OBJETIVOS

El objetivo principal de esta línea de trabajo es fortalecer el aprendizaje crítico y significativo de la Informática en la escuela, con especial énfasis en la enseñanza de la programación y el uso responsable de las TIC; atendiendo a la formación de los ciudadanos del siglo XXI que requieren de nuevas habilidades y destrezas vinculadas a diseñar, crear e innovar con los medios digitales y no ser consumidores pasivos de tecnologías digitales.

Se proponen los siguientes objetivos específicos:

- Elaborar materiales didácticos que permitan el trabajo de y con Informática en las aulas de las escuelas.
- Desarrollar herramientas didácticas que complementen a otras disciplinas, enriqueciendo la propuesta educativa y permitan trabajar la Informática en la escuela en todos los niveles (inicial, primaria y secundaria) aplicando conceptos de *gamification*¹ y juegos serios, entre otros.
- Generar un espacio virtual que concentre los materiales y herramientas desarrolladas, de libre acceso y disponibilidad.
- Diseñar e implementar intervenciones con

¹ El término *Gamification* hace referencia al uso del diseño y la mecánica de juegos para mejorar los contextos no relacionados con los juegos.

docentes y estudiantes de los diferentes ciclos escolares.

- Evaluar las intervenciones realizadas en los distintos contextos.

Enmarcadas en esta línea de investigación se han desarrollado herramientas tanto en el marco de tesis de posgrado, tesinas de grado como en trabajos de cátedra y proyectos propios del LINTI. Entre las herramientas desarrolladas se pueden mencionar a: RITA (Aybar Rosales, 2015), RITA en RED (Aybar Rosales, 2017), XRemoteBot (López, 2016), DROPSY² y DuinoBotSocks³. Actualmente algunas de ellas se encuentran en uso y forman parte del material didáctico con el que se trabaja en los espacios de formación docente y las actividades con estudiantes de escuelas; otras están en procesos de pruebas y evaluación. Asimismo se trabaja articuladamente con asignaturas de las carreras de la Facultad: en la materia “Seminario de Lenguaje-opción Python”, los estudiantes supervisados por sus docentes, desarrollan herramientas destinadas a la enseñanza de programación de niños/as del nivel inicial y primer año del nivel primario.

Durante el año 2017 se llevó a cabo una experiencia que permitió elaborar y poner en práctica una propuesta para enseñar programación en el nivel primario y secundario. Esta propuesta contempla una nueva herramienta didáctica para incorporar la programación de robots físicos (Harari Viviana y Banchoff Claudia, 2015). Esta intervención es el resultado de una tesis de maestría en “Tecnologías aplicadas a Educación”: “ProBots3D: una herramienta libre para enseñar programación a niños y jóvenes”, actualmente en proceso de evaluación.

En las actividades de esta línea de investigación, se ha participado en dos proyectos impulsados por la Fundación Sadosky⁴ cuyo objetivo es aportar desde una visión federal a la incorporación de la disciplina Informática en las escuelas:

- La elaboración de un manual de Informática destinado a docentes del segundo ciclo de

nivel secundario, que consta de 12 capítulos; actualmente está en proceso editorial con fecha prevista de publicación durante el primer semestre de 2018.

- El diseño curricular y dictado de la primera cohorte de una “Especialización docente en didáctica de las Ciencias de la Computación” junto con el ISFD N° 95 de La Plata, destinada a docentes de nivel medio de la provincia de Buenos Aires.

Estos proyectos, dan un marco para indagar sobre el impacto de la incorporación de esta disciplina en las actividades escolares y trabajar en la formación de docentes.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está integrado por docentes-investigadores del LINTI y estudiantes de la Facultad de Informática, quienes han ido completando su formación tanto de grado como de postgrado a lo largo de estos años.

En este sentido, se han formulado varias tesinas, tesis de postgrado, proyectos de extensión y actividades de cátedras relacionadas con las herramientas involucradas. Actualmente se encuentran en desarrollo varias tesinas de grado y tesis de postgrado que contribuirán en esta línea de investigación.

5. BIBLIOGRAFÍA

Aybar Rosales Vanessa, Queiruga Claudia, Banchoff Tzancoff Claudia, Kimura Isabel Miyuki y Brown Bartneche Matías (2017). *Programming Competitions in High School Classrooms: RITA en RED*. En 2017 XLIII Latin American. Córdoba, 4 al 8 de septiembre de 2017. Editorial: IEEE. ISBN: 978-1-5386-3057-0. Indexada: DBLP, IEEE Xplore.

Aybar Rosales Vanessa, Queiruga Claudia, Kimura Isabel, Brown Barnetche Matías y Gómez Soledad (2015). *Enseñando a programar con RITA en escuelas secundarias*. En el proceeding de la XXI Congreso Argentino de Ciencias de la Computación (CACIC 2015), Junín, Argentina, Octubre 5-9, 2015. ISBN 978-987-3724-37-4, pp. 1201-1211.

Recuperado de:

² Disponible en <https://github.com/dropsy-unlp>

³ Disponible en <https://github.com/Robots-Linti/DuinoBotSocks>

⁴ Fundación Sadosky:
<http://www.fundacionsadosky.org.ar/>

<http://sedici.unlp.edu.ar/handle/10915/50644>.

Díaz Javier, Banchoff Tzancoff Claudia, Queiruga Claudia y Martín Sofia (2014). *Experiencias de la Facultad de Informática en la Enseñanza de Programación en Escuelas con Software Libre*. En las Memorias del Congreso Iberoamericano de Ciencia, Tecnología, Innovación y Educación 2014, Buenos Aires, Argentina, Noviembre 12-14. ISBN 978-84-7666-210-6. Artículo 1426. Recuperado de: <http://www.oei.es/congreso2014/memoriactei/1426.pdf>

Dirección General de Cultura y Educación de la Provincia de Buenos Aires (2018). *Diseño curricular para la educación primaria: primer ciclo y segundo ciclo*; coordinación general de Sergio Siciliano. - 1a ed. ISBN 978-987-676-095-9.

Recuperado de: <http://servicios.abc.gov.ar/lainstitucion/organismos/consejogeneral/disenioscurriculares/primaria/2018/dis-curricular-PBA-completo.pdf>

Dirección General de Cultura y Educación de la provincia de Buenos Aires (2010). *Diseño Curricular para la Educación Secundaria Ciclo Superior. ES4: Nuevas Tecnologías de la Información y la Conectividad*. Coordinado por Claudia Bracchi. -1a ed. ISBN 978-987-1266-98-2.

Recuperado de: http://servicios.abc.gov.ar/lainstitucion/organismos/consejogeneral/disenioscurriculares/secundaria/materias_comunes_a_todas_las_orientaciones_de_4anio/tic_4_final_web.pdf

Harari Viviana y Banchoff Claudia (2015). *La Facultad en 3D: un entorno común para dos propuestas educativas*. En libro de actas del XXI Congreso Argentino de Ciencias de la Computación (CACIC 2015). Octubre, Junín, Argentina. ISBN 978-987-3724-37-4.

López Fernando; Banchoff Tzancoff Claudia y Queiruga Claudia (2016). *XRemoteBot*. II Jornadas Argentinas de Tecnología, Innovación y Creatividad. Universidad CAECE, sede Mar del Plata; 2, 3 y 4 de Noviembre. En Malbernat L. R.; Finochietto J.R.; Bacigalup G.F. (Comps). ISBN 978-987-46267-0-7.

National curriculum in England: computing programmes of study (2013). Recuperado de: <https://www.gov.uk/government/publications/national-curriculum-in-england-computing-programmes-of-study>.

PLANIED: *Plan Integral de Educación Digital del Ministerio de Educación y Deportes de la Nación Argentina* (s.f). Recuperado de: <http://planied.educ.ar/category/novedades/>

Program.AR: programa de la Fundación Sadosky, Ministerio de Ciencia, Tecnología e Innovación Productiva de la Nación Argentina (s.f).

Recuperado de: <http://www.fundacionsadosky.org.ar/programas/programar/>

Queiruga Claudia, Banchoff Tzancoff Claudia, Martín Sofia, Aybar Rosales Vanessa, López Fernando, Kimura Miyuki Isabel y Soledad Gómez (2017). *PROGRAMAR en la Escuela: Nuevos Desafíos en las Aulas*. En XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017), Ciudad de Buenos Aires, 27 y 28 de Abril. ISBN 978-987-42-5143-5, pp 732-736. Editorial: Red de Universidades con Carreras en Informática (RedUNCI).

The New Zealand Curriculum on-line (s.f). Disponible: <http://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum/Technology>

Wing, Jannette M. (2006). *Computational Thinking*. Communications of the ACM, vol. 49, 33-35.

Wing, Jannette M. (2008). *Computational thinking and thinking about computing*. Philosophical Transactions of The Royal Society A, vol. 366, 3717–3725.

La tecnología como vehículo de articulación Nivel Medio / Universidad
 Fernanda Beatriz Carmona, Alberto Eduardo Riba, Fernando Emmanuel Frati,
 Claudia Isaia, Alejandro Cruz, Jorge Tejada, Matías Pérez, Patricia
 Manriques, Emmanuel Alejandro Portugal Murcia, José Nicolás Frati, Pablo
 Oporto Poblete, Sergio Arévalo

Departamento de Básicas y Aplicadas, Universidad Nacional de Chilecito
 9 de Julio 22, Chilecito, La Rioja, Argentina
 {fbcarmona, ariba, fefrati, cisaia, acruz, jtejada}@undec.edu.ar, {mataguper20,
 patriciamanriques311995, emmanuel.portugal.91, nicofrati, pablogabrieloporto,
 sergioa}@gmail.com

Resumen

Esta línea de I+D pretende fortalecer la articulación Universidad – Nivel Medio, estimulando el diseño de mecanismos que permitan alcanzar un diagnóstico compartido y planes de trabajo, destinados a disminuir la brecha que existe entre estos dos niveles.

Se propone una estrategia pedagógica / tecnológica como soporte de los procesos de enseñanza y aprendizaje a ser utilizados en las Instituciones Educativas (IE) de nivel medio para despertar vocaciones tempranas vinculadas a las carreras tecnológicas, utilizando la Robótica Educativa, el aprendizaje colaborativo y por descubrimiento guiado y el estímulo basado en competencias.

Los temas abordados son transversales a varias áreas, como educación, algoritmos, lógica, programación y arquitectura de computadoras.

El desarrollo de esta línea conlleva, sin dudas, a la generación de otros proyectos relacionados con el uso de la tecnologías en la articulación Universidad - Nivel Medio considerando la amplia cobertura de la temática a desarrollar y la imperiosa necesidad de articulación existente entre estos niveles educativos, como así también, en la gestación de líneas de investigación relacionadas con nuevas estrategias didácticas en Robótica y Educación, la enseñanza y aprendizaje de

la programación, la utilización y programación de microcontroladores en otros ambientes de trabajo y la utilización de lenguajes interpretados.

Palabras clave: Educación, capacitación, TIC, enseñanza-aprendizaje, robótica educativa, aprendizaje experimental, aprendizaje inductivo, programación, juego.

Contexto

Esta línea de I/D/I corresponde al desarrollo e implementación de proyectos que fortalecerán la inserción de UNDeC en la comunidad y especialmente su articulación con los demás niveles educativos. Refiere al proyecto “La robótica como introducción a la formación tecnológica” SPU 2013-2015 y la colaboración de otras instituciones del país y del extranjero a través de los proyectos “Red para la Integración de Universidades en el uso de TIC para la Inclusión en la Educación Superior” aprobado en la VII Convocatoria a Redes Internacionales, año 2013, el proyecto “Red para la creación y publicación de objetos virtuales de aprendizajes de calidad en Repositorios Institucionales” aprobado en la IX Convocatoria a Redes Internacionales, 2016-2017, el proyecto “Mejora de la Enseñanza de las Ciencias” convocatoria La Universidad y la Escuela Secundaria, tercera etapa 2015-2017,

SPU – Ministerio de Educación de la Nación, el Proyecto “Repositorios Digitales con Contenidos Orientados a las Necesidades de Escuelas Rurales (ER)” por Resolución CE N° 1055/15 convocatoria de Proyectos de Desarrollo Tecnológico y Social (PDTS), 2016-2018 y el proyecto "Programación Colaborativa" resolución N° EXP-S01:0001042/2016, período 2016-2017 "Universidad, Cultura y Sociedad", SPU y su continuidad 2017-2018 aprobada por RESOL-2017-5135-APN-SECPU.

Introducción

Con la constante evolución de las Tecnologías de la Información y la Comunicación (TIC) se hace necesaria la incorporación, integración y aprovechamiento pedagógico de éstas en el Sistema Educativo. Por otro lado, vivimos una época conocida como la “Era de la Información”, caracterizada por una gran demanda a nivel nacional e internacional por profesionales calificados en el uso de TIC [1, 2,3]. En la Argentina esta realidad ha sido identificada como clave para la innovación, incremento de las exportaciones, actividades de investigación, emprendedurismo y producción de valor agregado para todos los sectores de la sociedad. Evidencia de ello son la existencia de distintas políticas de estado que promueven el desarrollo del sector, como ser la Ley de Software (Ley N° 26692), becas para jóvenes profesionales TIC [4] o la puesta en funcionamiento de la iniciativa Program.AR [5], que promueve la enseñanza de la computación en todas las escuelas argentinas. La implementación del programa del Estado Nacional denominado Modelo 1 a 1 a través del programa “Conectar Igualdad” al cual la provincia de La Rioja complementó con el plan “Joaquín V. González”, han

alcanzado una alta cobertura, cumpliendo con la entrega de netbooks a estudiantes y docentes de escuelas secundarias. Sea cual sea el nivel de integración de las TIC en los centros educativos, los docentes y estudiantes necesitan una “alfabetización digital” y una actualización didáctica.

Es necesario entonces fortalecer la articulación Universidad / Nivel Medio, planteando nuevos escenarios donde los estudiantes no solo sean usuarios de las tecnologías sino participantes activos, generando una adecuada planificación y realización conjunta de acciones que favorezcan la inserción de los estudiantes en la Universidad para prevenir y disminuir las causas de deserción y estancamiento en los primeros años de las carreras universitarias. Esta línea aborda dos aristas del mismo problema, la escasa matrícula de ingresantes que optan por una carrera TIC y el marcado nivel de deserción en los primeros años de estudio de quienes lo hicieron.

La robótica y la programación como introducción a la formación tecnológica

Los procesos de desarticulación y segmentación educativa, desarrollados en los últimos años en nuestro país y en especial en la región, dificultan el pasaje, ingreso y permanencia, de los alumnos del nivel medio a los estudios universitarios [6]. Las instituciones universitarias requieren de sus alumnos aprendizajes vinculados a las destrezas y estilos de pensamiento de las culturas disciplinares, y el desarrollo de una personalidad autónoma y crítica para su desempeño social [7].

Pese a la marcada demanda por profesionales en carreras TIC, es notable que el número de estudiantes que eligen estas carreras sea muy inferior al de estudiantes que eligen carreras más clásicas. Aunque existen múltiples motivos que pueden explicar esta situación (vocación, popularidad de las

carreras, expectativas profesionales, entre otras) se destaca particularmente algunos “prejuicios” asociados a las carreras TIC: son carreras exclusivas para jóvenes sobresalientes (particularmente en matemáticas), la tarea de programar es difícil, aburrida y/o repetitiva, hay que saber mucho inglés, entre otras.

Con el fin de acortar la brecha existente en la articulación Universidad / Nivel Medio se deben profundizar contenidos teóricos - prácticos y su aplicación en un proceso de formación integral que reúna conocimientos, habilidades, destrezas y actitudes propios de los modos de producción en los diferentes campos disciplinares, acordes a los requerimientos sociales.

Conforme a estas tendencias, consideramos oportuna la intervención de la UNDeC como nexo de articulación entre Nivel Medio y Universidad, dado que la universidad es el medio adecuado y pertinente para fortalecer los procesos de enseñanza y aprendizaje sobre dicha tecnología, con el objetivo final de despertar en los estudiantes de Nivel Medio la vocación por las carreras tecnológicas, particularmente en Informática.

La Robótica Educativa se concibe como un contexto de aprendizaje que se apoya en las tecnologías digitales y en los procesos de mediación pedagógica para que los estudiantes creen prototipos o simulaciones robóticas que surgen a partir del ingenio, la creatividad y puesta en práctica de lo aprendido [8]. Es de interés plantear a la robótica como vehículo de aprendizaje con materiales concretos, motivando a los jóvenes a construir, diseñar y explorar nuevas formas de hacer las cosas, a través del aprendizaje experimental, el trabajo en equipo y el desarrollo de su confianza y habilidades innovadoras, brindando un espacio que les permita no sólo ser usuarios de las

tecnologías, sino que, a partir de conocimientos matemáticos, mecánicos, físicos y lógicos, logren resolver, en forma activa, problemas significativos.

En Argentina y en Latinoamérica se están implementando proyectos, como propuestas de enseñanza de la programación en los primeros años de la carrera, que incorporan entre sus estrategias el aprendizaje basado en problemas y el aprendizaje colaborativo, otros incluyen, además, a la robótica como una opción para la profundización y gestación de habilidades cognitivo-creativas [10, 11, 12].

Algunas de estas propuestas promueven la construcción de robots que compiten de acuerdo a reglas internacionales y por categorías, otras usan la robótica como recurso de apoyo en el estudio de habilidades básicas en matemáticas, ciencias o física y construyen y programan modelos que ayudan a representar con elementos externos esos conceptos. Otros promueven la construcción de robots que ejecutan tareas y funciones particulares o que se comportan de cierta manera ante variables del ambiente [13, 14].

Aunque todas ellas sirven de inspiración, aplicamos un enfoque diferente: utilizar actividades colaborativas y lúdicas relacionadas con la programación de un robot para la resolución de problemas, como estrategia de enseñanza de la programación destinada especialmente a estudiantes del Nivel Medio [15].

No buscamos crear un curso completo de programación, sino generar un espacio de acercamiento a la tecnología y al mundo de la programación que resulte atractivo para los jóvenes antes de que ingresen a la universidad. Por otro lado la experiencia adquirida en los últimos años como organizadores de la sede Chilecito del Torneo Argentino de Programación (TAP) nos ha permitido poner en valor

para el aprendizaje de la programación la motivación obtenida por las competencias y el reconocimiento entre pares: aquellos estudiantes que han participado del torneo han mejorado su desempeño académico, se muestran más dispuestos a compartir lo que saben y participan con más frecuencia en las propuestas académicas extracurriculares.

Se desarrolló una comunidad virtual de aprendizaje colaborativo de la programación. El soporte tecnológico para esta comunidad está dado por una red social diseñada y desarrollada por el equipo de trabajo y orientada a fomentar el aprendizaje de la programación entre pares. Los miembros de esta comunidad en lugar de ocupar jerarquías formales establecidas por los roles de los participantes, adquieren una “reputación” dentro de la comunidad basada en distintas actividades realizadas a través de la plataforma. La reputación de los miembros permite construir un ranking similar al que se tiene en cualquier juego basado en la competencia. Las actividades para adquirir créditos o aumentar la reputación consisten en resolver problemas de programación usando distintos lenguajes, proponer nuevos problemas, ayudar a otro miembro con dificultades o describir soluciones y técnicas de programación. Se busca fomentar el aspecto lúdico de la competencia donde la participación es de carácter individual, complementado con la cooperación o colaboración entre pares, donde se premia el desarrollo de habilidades grupales.

Consideramos importante complementar ambas propuestas permitiendo interactuar jóvenes de los dos niveles educativos con la finalidad de reducir la deserción en los primeros años de la carrera.

Líneas de Investigación, Desarrollo e Innovación

- Impacto de las TIC en el proceso de enseñanza-aprendizaje del Nivel Medio.
- Desarrollo de Objetos de Aprendizaje.
- Articulación Universidad - Nivel Medio.
- Programación colaborativa.
- Robótica Educativa.

Resultados y Objetivos

El desarrollo de esta línea permitirá:

- Fortalecer la relación entre Universidad e IE del Nivel Medio.
- Promover el desarrollo de habilidades de resolución de problemas de programación, trabajo colaborativo y vinculación entre estudiantes de nivel secundario y estudiantes de nivel universitario de carreras TIC.
- Desmitificar los prejuicios clásicos asociados a la programación entre los estudiantes de nivel secundario y universitario en carreras TIC.
- Generar un ambiente de aprendizaje que permita a los alumnos integrar distintas áreas del conocimiento, adquiriendo habilidades generales y nociones científicas.
- Constituir un equipo interdisciplinario para la investigación y desarrollo de contenidos educativos y estrategias didácticas en TIC, Robótica y Educación.
- Estimular las prácticas de estudio colaborativo entre pares, la competencia sana y el reconocimiento del saber académico.
- La expansión de las TIC en las Escuelas de Nivel Medio del departamento Chilecito y de la región.
- Disminuir del índice de deserción en el primer año de las carreras incluidas como oferta académica de la UNDeC.

Actualmente se ha diseñado y construido el kit de Robótica Educativa que está compuesto por un módulo principal que

aloja al microcontrolador Arduino [9] UNO R3, un sistema de alimentación, un conjunto de sensores que componen el sistema de percepción del robot y actuadores que permiten modificar las variables del entorno controladas. Para su construcción se utilizaron piezas de bajo costo y disponibles en el mercado nacional para que en caso de pérdida o daño puedan ser reemplazadas fácilmente. Se incluye una librería con funciones, programadas por el equipo de trabajo, que se anexan al entorno de desarrollo Arduino, con el objeto de brindarle al estudiante un nivel de abstracción superior que le permita programar en un lenguaje más natural e intuitivo que el aportado por el propio entorno. Se adquirió una impresora 3D con la cual se diseñan y construyen los distintos componentes del kit de Robótica Educativa.

Durante el año 2015 y 2017, se desarrollaron talleres con la participación de estudiantes de IE de nivel medio. Cada taller está constituido por tres módulos, un módulo permite que los estudiantes programen cada uno los componentes del kit experimentando el comportamiento de estos antes los estímulos incluidos, otro módulo muestra el funcionamiento de la impresora 3D y las aplicaciones para el diseño de las diferentes piezas, y con el tercero los estudiantes experimentan con el robot programando la resolución de un juego. En grupo de a diez, los estudiantes recorren cada módulo.

Con el desarrollo de la primera etapa de ejecución del proyecto de Programación Colaborativa se adquirió un servidor para alojar la plataforma de programación colaborativa. El servidor se encuentra online y aloja la plataforma de programación. Esta plataforma se desarrolló a partir de la integración de dos herramientas opensource disponibles, ELGG y BOCA. ELGG es un framework

para desarrollar sistemas con requerimientos de redes sociales; BOCA es el software utilizado durante las competencias de programación del TAP y de la ICPC-ACM. La primera versión se utilizó durante el desarrollo del taller de programación competitiva (agosto 2017) y se mantuvo online hasta la competencia internacional ICPC-ACM (noviembre 2017) en el sitio <https://etic.undec.edu.ar/code>, llegando a tener más de 50 usuarios activos. Actualmente se está trabajando en una versión mejorada de la plataforma desplegada en el mismo sitio. Se realizó un Taller de Programación Competitiva con la asistencia de 30 jóvenes estudiantes de la UNdeC y de un secundario técnico con orientación en informática de la región. La mayoría de los asistentes participaron en el mes de septiembre de las competencias de programación TAP (nivel universitario) y Olimpiada Informática Argentina OIA (nivel medio).

Avances de la línea fueron expuestos en el artículo "Robótica educativa: una estrategia para despertar vocaciones tempranas en Informática." aprobado y presentado en el IX Congreso sobre Tecnología en Educación y Educación en Tecnología, Chilecito (La Rioja) 2014; en la Sesión de Demos Educativos en el X Congreso de Tecnología en Educación y Educación en Tecnología, Corrientes 2015 y en las III, IV, V y VI Jornadas Científicas de Estudiantes Investigadores UNdeC, 2014 / 2015 / 2016 /2017 con la participación de alumnos de grado.

Formación de Recursos Humanos

El equipo de investigación de esta línea de trabajo está compuesto por seis docentes de la Universidad Nacional de Chilecito y seis alumnos de grado. De los docentes: uno es doctor en ciencias de la computación especializado en cómputo

paralelo y tecnología grid, dos docentes que se encuentran desarrollando su tesis de Maestría en Informática uno en el área de Objetos de Aprendizaje y el otro en Mejora del Posicionamiento de Satélites y dos auxiliares docentes egresados de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas de la UNDeC. También participan 2 alumnos avanzados de grado que se encuentran desarrollando su trabajo final para la obtención del título de Ingeniero en Sistemas.

Los integrantes son docentes de las asignaturas Algoritmo y Estructuras de Datos, Arquitecturas de Computadoras, Arquitecturas Paralelas, Programación.

Referencias

[1] NIDIVERSIDAD. Las TIC, su importancia en la actualidad y el mercado laboral, 2016.

<http://www.universidad.com.ar/las-tic-su-importancia-en-la-actualidad-y-el-mercado-laboral>

[2] El Mundo. Faltan 900.000 profesionales TIC | Economía | EL MUNDO. (s. f.), 2016.

<http://www.elmundo.es/economia/2015/05/31/5568a4a1268e3e9e518b4592.html>

[3] El Espectador. «Hay déficit de 15.000 ingenieros» para industria TIC: Mintic, 2014.

<http://www.elespectador.com/noticias/economia/hay-deficit-de-15000-ingenieros-industria-tic-mintic-articulo-503625>

[4] MINCYT. Apertura de Becas Jóvenes Profesionales TIC 2016.

<http://www.agencia.mincyt.gob.ar/frontend/agencia/post/2311>

[5] Program.AR | Acercando a los chicos a la computación. (s. f.).

<http://program.ar/>

[6] Josfíe Joaquín Brunner and Rocío Ferrada Hurtado, “Educación superior en Iberoamérica”, Eds. Santiago, Chile: RIL R editores, Oct. 2011.

[7] SPU, “Articulación escuela secundaria educación superior,” Blog: Portal de Educación, 2007.

[8] Ana Lourdes Acuña, Marfía Dolores Castro, and Diana Matarrita Obando, “Desarrollo de capacidades para el diseño e implementación de proyectos de robótica educativa en América Latina y el Caribe,” Fondo. Reg. para la Innov. Digital en América Latina y el Caribe y la Fund. Omar Dengo, Informe Final de investigación, 2011.

[9] Massimo Banzi, “Getting Started with Arduino”, 2nd ed. O’Reilly Media, 2011.

[10] A. H. González and M. C. Madoz, “Utilización de TIC para el desarrollo de actividades colaborativas para la enseñanza de la programación,” Jul. 2013.

[11] R. Coppo, J. Iparraguirre, G. Feres, G. Ursua, and A. Cavallo, “Sistema didáctico para la enseñanza de la programación con metodologías de aprendizaje basado en problemas,” 2011, eje: Tecnología informática aplicada en educación.

[12] L. C. De Giusti, F. Leibovich, M. Sanchez, F. Chichizola, M. Naiouf, and A. E. De Giusti, “Desafíos y herramientas para la enseñanza temprana de concurrencia y paralelismo,” Oct. 2013, WIEI - II Workshop de innovación en educación en informática.

[13] Cristian Rigano and Juan Vivanco, “MHO: un robot de sumo - la primera experiencia en robótica con un LOGO!” Grupo de Robótica y Simulación, Departamento de Ingeniería Eléctrica, UTN-FRBB, Bahía Blanca, Reporte, 2006.

[14] M. A. Junco Rey, R. Swain Oropeza, A. Aceves López, and J. Ramírez Uresti, “RoboCup: el reto tecnológico de monterrey campus estado de México,” Arequipa, Perú, 2002.

[15] Fernanda B. Carmona, et al. “Robótica educativa: una estrategia para despertar vocaciones tempranas en Informática.” IX Congreso sobre Tecnología en Educación & Educación en Tecnología. La Rioja, 2014.

Edimbrujo: Hacia la definición de un modelo didáctico para la enseñanza de la Inteligencia Artificial en Juegos

Federico Amigone Pablo Kogan Gastón Michelan
Jorge Rodriguez

{fede.amigone, pablo.kogan,gaston.michelan,j.rodrig }@fi.uncoma.edu.ar

Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

Resumen

La amplitud de los campos disciplinares intervinientes en la enseñanza de las Ciencias de la Computación presenta serios desafíos para quienes intenten construir modelos didácticos para su enseñanza.

Una de las áreas de conocimiento que mayor complejidad reviste en la percepción general es la enseñanza de la Inteligencia Artificial, disciplina entendida como el diseño y análisis de agentes autónomos que perciben el entorno e interactúan de forma racional con él.

Por tratarse de un área emergente en materia de avances y estratégica respecto a sus implicancias futuras, resulta cada vez más necesario considerar su incorporar en los planes de estudio. Las fronteras de su aplicación se encuentran en permanente expansión y resulta necesario sostener una correspondencia con dicha progresión desde la dimensión curricular.

Este trabajo presenta una Línea de Investigación tendiente a consolidar una estrategia de abordaje para la enseñanza del núcleo disciplinar de la Inteligencia Artificial desde un enfoque lúdico. A tal efecto se describe un ecosistema digital construido como plataforma de soporte al proceso de enseñanza y de aprendizaje de las técnicas que sustentan el razonamiento artificial, siempre en el contex-

to de su aplicación a los juegos.

Palabras Clave: ENSEÑANZA DE LAS CIENCIAS DE LA COMPUTACIÓN, AGENTES, INTELIGENCIA ARTIFICIAL.

Contexto

Esta propuesta se ubica en el contexto de las iniciativas promovidas por el Grupo de Investigación en Lenguajes e Inteligencia Artificial de la Facultad de Informática.

Este trabajo se desarrolla en el ámbito del proyecto de investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación (04/F015)* que está financiado por la Universidad Nacional del Comahue a través de la Secretaría de Ciencia y Técnica. El proyecto tiene prevista una duración de cuatro años a partir de enero del 2017.

1. Introducción

La aplicación de técnicas de Inteligencia Artificial se encuentra presente en muchos aspectos de la interacción de los sujetos con el medio social y económico. Esta situación contribuye a su consideración, como área de conocimiento, en documentos de recomendaciones curriculares para carreras de Informática [9, 8].

Esta línea parte del supuesto de que el campo disciplinar en referencia presenta grandes oportunidades para construir modelos didácticos basados en la construcción de procesos de enseñanza más lúdicos y motivadores con excelentes oportunidades para plantear dinámicas grupales y favorecer la retención del interés del estudiante[2].

La Línea de Investigación y Desarrollo propone el diseño de una plataforma para la enseñanza de las Ciencias de la Computación, en particular, en el contexto del área de conocimiento Inteligencia Artificial: Edimbrujo.

El desafío requiere desarrollar dos productos disímiles que actúan en conjunto para favorecer el proceso de aprendizaje de conceptos y prácticas relacionados a la Inteligencia Artificial.

Por un lado, el modelo didáctico basado en la capacidad con que cuentan los ambientes lúdicos para mejorar las posibilidades de aprender conceptos y prácticas sobre Inteligencia Artificial[3, 1]. Por otro, un framework tecnológico que permita a los estudiantes construir y disfrutar de una experiencia de laboratorio orientada al desarrollo de un artefacto computacional lúdico haciendo uso de técnicas de Inteligencia Artificial[2].

La mejora de la motivación y retención del interés por parte del sujeto de aprendizaje contribuye a aumentar el rendimiento académico y sus habilidades personales como resultado del aprendizaje a través del juego[5]

El desafío no es menor, ya que la conformación de un ecosistema lúdico en conjunción con técnicas de Inteligencia Artificial requiere articular una arquitectura heterogénea tanto en sus componentes fundantes como en la tecnología utilizada.

Respecto del framework tecnológico sobre el cual opera el modelo, es importante notar que pese a la evidente complejidad que puede representar convocar al uso tecnologías heterogéneas, el motor emocional que tracciona el proceso de aprendizaje en el contexto de la construcción de juegos de competencia emerge como un motor motivacional suficiente para llegar a buen término. Es decir, el modelo

didáctico debe actuar como el propulsor lúdico para vehicular el proceso a través de una curva de aprendizaje leve.

2. Línea de investigación y desarrollo

Edimbrujo es una plataforma sobre la que es posible programar Agentes Inteligentes que interactúen en un contexto lúdico, cuyo comportamiento se encuentra restringido al Mundo o Historia asociado a Edimbrujo. Sobre ella opera un modelo Didáctico basado en la colaboración grupal para la competencia entre pares cuya finalidad persigue la enseñanza de técnicas de Inteligencia Artificial utilizando la programación de Agentes Inteligentes[1].

En esta sección se describen el modelo didáctico, los componentes arquitecturales que posibilitan el funcionamiento de Edimbrujo, las herramientas de diseño que permiten construir los escenarios y personajes y finalmente nos centraremos en el objetivo primario que es la experiencia educativa de desarrollar Agentes Inteligentes.

2.1. La Historia

Edimbrujo es tanto un juego como la plataforma tecnológica que lo soporta. Ahora bien, todo juego tiene una historia asociada. En la historia o relato, se define lo que en el argot lúdico suele denominarse Mundo o Historia, esto es, la contextualización de los personajes y su misión en la historia del juego[6].

En el caso de Edimbrujo la Historia podría resumirse así: Con el debacle del Imperio Romano las ciudades más lejanas quedan totalmente abandonadas, algunas esconden antiguos tesoros. La más famosa es Edimbrujo, en el fin del mundo donde ya no hay más tierra y solo mar, se encuentra la ciudad maldita, una ciudad donde habitan todo tipo de seres crapulentos, no obstante los antiguos tesoros ocultos llaman la atención de aventureros con sed de fama o riquezas. Se rumorea

que una vez al año, para el 31 de octubre es cuando la ciudad cobra vida y las más grandes riquezas afloran.

2.2. El Modelo Didáctico

El modelo de aprendizaje propuesto se basa fundamentalmente en un proceso lúdico grupal basado en competencias entre pares sobre la plataforma Edimbrujo.

A tal efecto, se organiza primeramente una instancia de definición de equipos y se propone construir una identidad social que fortalezca el sentido colectivo y la identificación con el par: el grupo debe tener un nombre y estrategias de supervivencia.

Luego se propone una instancia creativa en el marco de Edimbrujo. Esta es una oportunidad para que el grupo genere una dinámica creativa y desestructurada, situación que en líneas generales los pone del lado contrario al rol que suelen tener en su interacción con los juegos: de experimentadores pasan a diseñadores. Así que se estimula la producción creativa de personajes, nombres, intereses, objetivos, peligros y dificultades.

En cualquier caso, el modelo alienta instancias en las que cada estudiante pueda valorar su participación como co-creador de una historia, a través de la generación de aportes del mundo Edimbrujo.

Luego de la participación creativa, los grupos deberán implementar Agentes Inteligentes que insuflen vida a los nuevos personajes diseñados haciendo uso de técnicas de Inteligencia Artificial y respetando las reglas del mundo Edimbrujo.

Una vez finalizada la implementación de las estrategias y técnicas de los Agentes y fundamentados los criterios y decisiones por parte de los grupos desarrolladores, tiene lugar La Competencia.

2.3. Componentes de la Plataforma Edimbrujo

Edimbrujo posee una arquitectura cliente servidor que hace uso de determinados com-

ponentes y protocolos.

2.3.1. Motor Edimbrujo

Para el armado del componente arquitectural que hace las veces de servidor o motor de Edimbrujo, se utiliza el lenguaje JavaScript sobre NodeJs. El motor agrupa la lógica del juego, el servidor socket.io encargado de la comunicación con clientes web y el servidor REST que se ocupa de la comunicación con agentes autónomos.

2.3.2. Clientes Edimbrujo

Edimbrujo es accesible a usuarios finales, es decir jugadores, través de un cliente web. Pero también resulta necesario un mecanismo de conexión entre los Agentes Autónomos y el motor Edimbrujo. Para ello se utilizan los servicios web REST publicados por el motor y que son consumidos por los procesos Agentes Autónomos desarrollados en php, python y/o java.

2.3.3. Escenarios Edimbrujo

Para la creación del mundo se utiliza el programa Tiled Map Editor, un editor libre y de código abierto que permite diseñar el mundo a través de capas de objetos, colocando directamente imágenes sobre un lienzo y luego exportando el resultado final como un archivo CSV o JSON, el cual será utilizado en Phaser para representar el mundo[10].

Phase es un framework open source, escrito íntegramente en Javascript, que se utiliza para construir juegos móviles o de escritorio en HTML5. Este framework provee una serie de herramientas que aceleran el desarrollo y ayudan a manejar las tareas genéricas necesarias para completar los juegos, como lo son la creación y manejo de escenarios, controles, movimientos y animaciones[7].

Para diseñar el mundo se utiliza una imagen compuesta por un conjunto de imágenes, todas del mismo tamaño. Cada imagen puede utilizarse para armar combinaciones simples o complejas del escenario.

Para completar la creación del escenario resulta necesario diseñar sprites o avatares de los personajes. Naturalmente el diseño que se logre debe tener su correspondiente correlato con el tipo de acciones que cada personaje podrá ejecutar.

2.3.4. Servicios REST

Se utiliza un pool de servicios REST como canal agnóstico de intercambio de datos entre el cliente autónomo y el motor Edimbrujo. Este protocolo de comunicación se basa en el desarrollo de HORNERO (Gestor de torneo de programación) [4]. Esto permite una interfaz muy débilmente acoplada entre los agentes y el motor, cualidad importante a la hora de desarrollar Agentes en cualquier lenguaje.

2.3.5. Comunicación con el cliente

WebSockets es una tecnología que se utiliza para crear canales de comunicación bidireccional en aplicaciones cliente-servidor, permitiendo así el intercambio de información entre estas dos entidades a través de un único socket TCP.

La tecnología WebSocket es un estándar de HTML5 que se puede utilizar en la mayoría de los servidores web, entre ellos Apache y Node.js. El protocolo WebSocket se encuentra definido en el documento RFC 6455.

La implementación de WebSockets varía dependiendo del lenguaje de programación que se utiliza, principalmente en el servidor, sin embargo la estructura de la aplicación es la misma. El objetivo es crear un WebSocket en el cliente y utilizarlo para enviar datos al motor Edimbrujo. La estructura de los datos se define a partir de un lenguaje de interfaz, generalmente se utiliza JSON debido a que en el cliente los WebSockets se implementan mediante Javascript.

En el servidor se deben implementar las funciones que procesan la información enviada desde el cliente a través del WebSocket. Como el canal de comunicación es bidireccional el servidor podría enviarle información al

cliente sin la necesidad de que exista un request por parte de este último, esta característica define a WebSocket como una herramienta muy importante para implementar juegos multiplayer.

En Edimbrujo los WebSockets se implementan mediante la librería Javascript socket.io en un ambiente basado en Node.js. Esta librería posee las funciones necesarias para la transferencia de datos entre el cliente y el servidor en tiempo real.

2.4. El Agente Autónomo

Llegamos así al componente más valioso en términos didácticos para el modelo propuesto sobre Edimbrujo. El Agente debe ser desarrollado por los grupos en PHP, Python o Java y debe contar con dos partes: Comunicación REST con Edimbrujo y Estrategia

La Estrategia es la dimensión del Agente donde se aplican las técnicas de Inteligencia Artificial, razón por la cual, debe comportar la mayor atención y esfuerzos de generación. A tal efecto, cada grupo debe defender los criterios y técnicas que utilizaron.

En general, el Agente desarrollado decide que acción realizar a partir del estado actual generando cada uno de los posibles estados próximos y eligiendo aquella acción que lo lleve al mejor estado. Para el caso de Edimbrujo, este estado se determina a partir de la distancia entre la próxima posición del jugador y un bono y/o enemigo más cercano y la vida actual del jugador. Este procedimiento se basa en el algoritmo de búsqueda local Hill Climbing. Al producirse un evento, tal como el de encontrar un enemigo o un premio, se dispara el análisis de reglas y las consideraciones de supervivencia del Agente.

3. Resultados

La primer experiencia con la plataforma Edimbrujo se realizó en el marco de la cursada de la materia optativa Inteligencia Artificial en Juegos, de la Licenciatura en Ciencias de la Computación de la Facultad de In-

formática de la Universidad Nacional del Comahue durante el segundo cuatrimestre del año 2017.

El enfoque propuesto hizo uso de Edimbrujo para instrumentar una instancia de competencia grupal en la que las estrategias de desarrollo del juego fueran conducidas por la ejecución de Agentes programados por los grupos de estudiantes.

La experiencia es evaluada positivamente ya que los objetivos perseguidos se cumplieron satisfactoriamente, es decir los estudiantes lograron construir aprendizajes sobre técnicas de Inteligencia Artificial en Juegos. Esto quedó experimentalmente constatado en el momento de la competencia, instancia en la que la programación de técnicas IA logró modelar comportamientos esperados en los agentes y se logró así una experiencia de aprendizaje lúdica de gran impacto.

4. Formación de Recursos Humanos

Se aspira a la formación como investigadores de los miembros más recientes del grupo. En este sentido se buscará fortalecer la vinculación con otros grupos de investigación e instituciones abocados a la temática en el ámbito nacional e internacional.

Asimismo, se espera la iniciación del posgrado de dos de los autores de este trabajo.

Referencias

- [1] C. C. Ada, T. M. Sanguino, S. Alacocer, A. Borrego, A. Isidro, A. Palanco, and J. Rodríguez. Classroom to mobile robots competition arena: An experience on artificial intelligence teaching.
- [2] J. Carpio Cañada, T. Mateo Sanguino, J. Merelo Guervós, and V. Rivas Santos. Open classroom: enhancing student achievement on artificial intelligence through an international online competition. *Journal of Computer Assisted Learning*, 31(1):14–31, 2015.
- [3] F. Chesani, A. Galassi, P. Mello, and G. Trisolini. A game-based competition as instrument for teaching artificial intelligence. In *Conference of the Italian Association for Artificial Intelligence*, pages 72–84. Springer, 2017.
- [4] C. C. Fracchia, P. Kogan, and S. Amaro. Competir + Motivar + Hornero = aprender programación. *Revista Iberoamericana de Educación en Tecnología y Tecnología en Educación*, 18:19–29, 2016.
- [5] D. Kirkland and F. O’Riordan. Games as an engaging teaching and learning technique: Learning or playing. In *1st Irish Conference on Engaging Pedagogies, Dublin, Ireland*, 2008.
- [6] K. L. McClarty, A. Orr, P. M. Frey, R. P. Dolan, V. Vassileva, and A. McVay. A literature review of gaming in education. *Gaming in education*, pages 1–35, 2012.
- [7] Phaser. <https://phaser.io/>.
- [8] Red UNCI. Documento de recomendaciones curriculares de la redunci 2014-2015, 2015.
- [9] M. Sahami, A. Danyluk, S. Fincher, K. Fisher, D. Grossman, E. Hawthorne, R. Katz, R. LeBlanc, D. Reed, S. Roach, et al. Computer science curricula 2013: Curriculum guidelines for undergraduate degree programs in computer science. *Association for Computing Machinery (ACM)-IEEE Computer Society*, 2013.
- [10] Tiled Map Editor. Home Page. <http://www.mapeditor.org/>.

METODOLOGÍA PARA LA DEFINICIÓN Y PONDERACIÓN DE FACTORES DE ÉXITO PARA PROCESOS DE GESTIÓN DE PROYECTOS ACADÉMICOS UNIPERSONALES DE PRÁCTICA PROFESIONAL SUPERVISADA EN CARRERAS DE INFORMÁTICA.

Rambo, A., Kuna, H., Sueldo, R., Urquijo, R., Piotroski F.

1. Depto. de Informática, Facultad de Ciencias Exactas Quím. y Naturales, Universidad Nacional de Misiones.

alirambo@fceqyn.unam.edu.ar

RESUMEN

Las carreras de formación profesional requieren como parte de la certificación, la realización de un proyecto que aborde de manera integral conocimientos adquiridos a lo largo de la formación recibida durante el cursado de la carrera seleccionada. Para ello se realiza un trabajo integrador, el cual tiene como objetivo dotar al estudiante de una experiencia en la práctica profesional cercana a su futuro laboral inmediato. Por tal motivo existe un espacio curricular en el cual los alumnos de manera individual y algunas veces grupal, deben realizar el análisis, diseño e implementación de un producto software, realizando en la última instancia la presentación y defensa del producto elaborado. Para tal fin el estudiante presenta una nota de solicitud de aprobación del tema elegido, esto es al principio del cursado de la cátedra abocada a su desarrollo, donde un equipo docente designado toma conocimiento del inicio del mismo, y a partir de ese instante acompaña al estudiante realizando un seguimiento y apoyo en las dudas que vayan surgiendo, cuidando de que se mantenga una evolución constante en el trabajo de manera tal que se finalice con el producto terminado en tiempo y forma.

La evaluación y seguimiento de los alumnos se realiza en proceso durante cada etapa de estos proyectos. Pero se detecta la necesidad de identificar los factores que determinan el éxito para los procesos de gestión de proyectos de desarrollo de software en el ámbito académicos contemplando las particularidades de este marco de trabajo.

Palabras clave: gestión de proyectos, unipersonales, práctica profesional supervisadas, sistemas de información, informática

CONTEXTO

Proyectos y Gestión de Proyectos de Software.

Un proyecto es “un esfuerzo temporal que se lleva a cabo para crear un producto, servicio o resultado único” según el PMI[1] quienes definen también a la gestión de proyectos como “la aplicación del conocimiento, habilidades, herramientas y técnicas a las actividades de un proyecto con la finalidad de conseguir los requisitos del mismo”.

La gestión de un proyecto involucra dentro entre las actividades a ser consideradas la de realizar actividades de gerenciamiento (definición, control, guía, monitoreo, selección, evaluación, etc.), planeamiento del proyecto, cronograma del proyecto, gestión de riesgos y estimación de costos con las particularidades que involucran a la gestión de proyectos de desarrollo de productos software [2].

Además las consecuencias de estimaciones inadecuadas desencadenan a posterior grandes pérdidas monetarias en los proyectos e incluso la no conclusión de los mismos [3].

SWEBOK1[4], Software Engineering Body of Knowledge, es un documento que fue creado por la Software Engineering Coordinating Committee, y se encuentra sustentado por la IEEE Computer Society2, se trata de una guía

¹ www.swebok.org

² IEEE Computer Society
www.computer.org/portal/web/guest/home

sobre el conocimiento presente en el área de la Ingeniería del Software.

Existen estándares de gestión de proyectos de tipo general, como ser el Project Management Body of Knowledge (PMBOK) del Project Management Institute (PMI)³ y también se encuentran disponibles enfoques de gestión específicos para proyectos de software. Actualmente se encuentra la 5ta edición disponible en español.

Entre los estándares específicos de gestión de proyectos de software se analizarán además de los mencionados anteriormente a los correspondientes al Capability Maturity Model Integrated (CMMI - capítulo Project Management)⁴ y al Rational Unified Process (RUP - también el capítulo de Project Management) por estar dichos enfoques entre los de mayor difusión [5][6]. Además existen otras normas y modelos que son seguidos por otras empresas, como CMM, ISO 9001, SPICE, PSP, TSP, ISO 20.000, ITIL, entre otros.

Factores de Éxito en Gestión de Proyectos.

El éxito de un proyecto debe medirse en términos de desviaciones del alcance de las características y funcionalidades previstas dentro de los márgenes de costos y tiempos previstos. Sin embargo los casos de retrasos, sobrecostos, frustración de expectativas e incluso fracasos rotundos en la consecución de los fines de los proyectos registrados en la literatura son muy significativos. En el campo de los proyectos de sistemas de información es muy referenciado el informe Chaos es una de las estadísticas más usadas. Son publicadas aproximadamente cada 2 años, desde 1994. El reporte CHAOS clasifica los proyectos en diferentes escalas que definen el *éxito* de los mismos cuando el proyecto fue entregado a tiempo, en el presupuesto y con todas sus funciones, se los califican como *deficiente* cuando el proyecto fue finalmente entregado pero con gastos muy superiores que están más allá del presupuesto, también aquellos que no se entregan a tiempo o no completados y como o *fracaso* cuando nada fue entregado. El último estudio del 2012 indica que el 39% de

todos los proyectos corresponden a casos exitosos, 43% se clasifican como deficientes y el 18% son casos que fracasaron [7].

Existen trabajos de diferentes autores que han intentado construir un marco para la clasificación de los factores críticos para el éxito o el fracaso de un proyecto. Según Pinto y Mantel[8] y Pinto y Prescott [9] se identifica que los factores críticos caen dentro de dos grandes grupos.

En el trabajo de Navascues [10] presenta los modelos y herramientas de la gestión de proyectos su empleo en relación con la simulación de procesos software para un entorno multiproyecto consigue identificar metodologías y modelos para descomponer jerárquicamente el problema multiproyecto, generar planes en condiciones de limitación de recursos y modelar y hacer frente al riesgo y la incertidumbre. El trabajo se valida sobre una empresa dedicada a realizar software a medida y busca optimizar la asignación de los recursos sobre los proyectos actuales para lo cual se procede a la validación sobre 50 proyectos pertenecientes a esta empresa.

En otro trabajo se presenta una metodología basada en los mapas cognitivos difusos para la formalización y el análisis de los factores críticos de éxito [11].

Hay estudios que buscan definir un modelo donde a partir de información básica del proyecto se pueda predecir su dificultad y clasificarlo en función de su riesgo [12]. Hay estudios llevados a cabo para analizar los factores de éxito y causas de fracaso más influyentes en proyectos donde aplican una primer etapa de análisis estadístico de encuestas realizadas y una segunda etapa de minería de datos y clusterización [13].

Minería de Datos.

La MD, en inglés Data Mining, se define como el proceso mediante el cual se extrae conocimiento comprensible, potencialmente útil, que previamente era desconocido de una BD, en diversos formatos y de forma automática [14].

Cabe destacar que la MD es una etapa dentro de un proceso más amplio que tiene como objetivo el descubrimiento de conocimiento en grandes BD (Bases de Datos) [15],[16], en

³ Project Management Institute www.pmi.org

⁴ Software Engineering Institute (CMU), Capability Maturity Model Integration www.sei.cmu.edu

inglés “Knowledge Discovery in Databases” (KDD).

Cuando se buscan antecedentes sobre el tema se ven trabajos sobre la aplicación de minería de datos para el análisis de métricas generadas en el desarrollo de proyectos de software en etapas previas como la especificación de requerimientos [17][18]. Por otra parte se encuentra un modelo para almacenar y recuperar métricas de software para realizar un seguimiento del proceso con respecto al tiempo, el costo y la calidad del mismo [19].

1 INTRODUCCION

1.1 Antecedentes

En otras Universidades con carreras de perfil de formación profesional similar también existen materias que abordan la ejecución de este tipo de proyectos entre ellas podemos mencionar las siguientes relevadas a modo de ejemplo: La materia Trabajo Final en el quinto año de la carrera de Licenciatura en Sistemas de cursado anual que se realiza en la Universidad Nacional de Chilecito, en cuyo fundamento menciona: “... Esta Asignatura contribuirá a formar un Ingeniero tecnológico capacitado para desarrollar sistemas de ingeniería y tecnología afines a los existentes y producir innovaciones. Formar un profesional capaz de analizar y evaluar requerimientos, y sobre esta base, desarrollar, diseñar, organizar e implementar sistemas de información...”⁵ En la Universidad Nacional del Nordeste en la cátedra Ingeniería del Software II, la cual se encuentra en el cuarto año de cursado de la misma, con la variante de que se tiene la opción de trabajar en grupos de hasta tres alumnos. En cuyos Objetivos menciona: “... Preparar al Alumno en la exposición de proyectos. Cubrir todas las etapas de documentación del proyecto. Desarrollo de habilidades de ‘Trabajar enfocado al cliente’; de esta manera el alumno deberá demostrar al cliente que el proyecto cumple las expectativas requeridas y acordadas. Uso eficaz del ciclo de vida del proyecto; utilizando el ciclo de vida que mejor se adapte para el mismo...”⁶ Por medio del relevamiento realizado se detecta que no existe una definición y ponderación de factores de éxito para procesos

de gestión de proyectos específicamente académicos, trabajar sobre la definición de los mismos puede suministrar la posibilidad de generar planes de acción con medidas de tipo preventivas para ser realizadas con los alumnos en desarrollo de estos proyectos previniendo el fracaso de los mismos.

2 LÍNEAS DE INVESTIGACION, DESARROLLO E INNOVACIÓN

Durante el desarrollo de este proyecto, la tesista abordara la investigación de los siguientes temas:

Planificación y gestión de proyectos de software

Factores considerados en los proyectos de software (énfasis en los relevados)

Explotación de la información y de su uso para evaluar factores de éxito en proyectos de ingeniería de software.

3 RESULTADOS Y OBJETIVOS

En el presente proyecto se prevé realizar:

Diseño de instrumentos y ejecución de los mismos orientados al relevamiento de datos vinculados a gestión de proyectos de software desarrollados en la universidad.

Diseño de instrumentos y ejecución de los mismos orientados a las tecnologías de explotación de información aplicables a la identificación de características presentes en proyectos de ingeniería de software que definan el éxito de los mismos.

4 FORMACION DE RECURSOS HUMANOS

Este proyecto es parte de las líneas de investigación del “Programa de Investigación en Computación” de la FCEQyN de la UNaM, con cuatro integrantes relacionados con las carreras de Ciencias de la Computación de la UNaM. De los cuales dos están realizando su tesis de pos-grado, uno se encuentran realizando tesis de grado.

5 BIBLIOGRAFIA

1. Project Management Institute. (2004). Guía de los fundamentos de la

5 http://www.undec.edu.ar/pdf/pdf_plan_es_LicenciaturaenSistemas.pdf
6 http://www.exa.unne.edu.ar/carreras/lic_sistemas_informacion.php

- Dirección de proyectos [Project Management Institute]. Project Management Institute, Newtown Square.
2. Sommerville, I., (2002) Ingeniería de software [trad de la 6ta edición], Addison Wesley, México.
 3. Wayt Gibbs W. (1994). La crisis crónica de la Programación. Scientific American.
 4. Bourque, P., Fairley, R. E. (2014). Guide to the Software Engineering Body of Knowledge (SWEBOK). IEEE Computer Society. Versión 3. Edition: 3.
 5. Jacobson, I., Booch, G., Rumbaugh, J., (2001) "Rational Unified Process" ISBN 0201707101, Addison Wesley.
 6. Equipo del Producto CMMI. (2012). CMMI para Desarrollo. Mejora de los procesos para el desarrollo de mejores productos y servicios. Software Engineering Process Management Program. Versión 1.3. , V1.3.
 7. Standish Corporation. (2003). Chaos chronicles version 3.0. Technical report, The Standish Group. <https://secure standishgroup.com/report s/reports.php>
 8. Pinto J. K., Mantel S. J. (1990) The causes of project failure. IEEE Transactions on Engineering Management, 37(4):269–276.
 9. Pinto J. K., Prescott J. E.(1990). Planning and tactical factors in the project implementation process. Journal of Management Studies 27(3):305–327.
 10. Navascués Fernández J. V., (2008) Técnicas avanzadas para la gestión de proyectos software. Universidad de Sevilla. Departamento de Lenguajes y Sistemas de Información. Trabajo para la obtención del Diploma de Estudios Avanzados. (Informe de Investigación). Postgrado Oficial en Ingeniería y Tecnología del Software <https://www.lsi.us.es/docs/doctorado/m emorias/Memoria-JNavascues.pdf>
 11. Vázquez L. M. Y., Rosello R. R., Estrada F. A. (2012) Modelado y análisis de los Factores Críticos de Éxito de los proyectos de software mediante Mapas Cognitivos Difusos. Ciencias de la Información, vol. 43, núm. 2, pp. 41-46. Instituto de Información Científica y Tecnológica. La Habana, Cuba. ISSN 0864-4659.
 12. Alba, C.; Rodríguez, V.; Ortega, F.; Villanueva, J. (2008). Predicción y clasificación de riesgos en proyectos de Sistemas de Información. 12th International Conference on Project Engineering. 9–11/07. Zaragoza. España.
 13. Cousillas, S. M.; Rodríguez Montequín, V.; Villanueva Balsera, J.; Alvarez Cabal, V. (2013). Project success Factors and Failure Causes Analysis: Behavioral Pattern Detection Using Clustering Techniques. 17th International Congress on Project Management and Engineering. 17-19/07. Logroño. España.
 14. Clark, P.; Boswell R. 2000. Practical Machine Learning Tools and Techniques with Java Implementation. Morgan Kaufmann Publisher.
 15. Fayyad, U.M., Piatetsky Shapiro, G., Smyth, P. (1996). From Data Mining to Knowledge Discovery: An Overview. (p 1-34). Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press.
 16. Britos, P., Grosser, H., Rodríguez, D., Garcia Martinez, R. (2008). [Detecting Unusual Changes of Users Consumption](#). (p. 297-306). In Artificial Intelligence and Practice II. Springer.
 17. Moreno García M. N., Quintales L. A., García Peñalvo F. J., Polo Martín M. J. (2001). Departamento de Informática y Automática Aplicación de técnicas de minería de datos en la construcción y validación de modelos predictivos y asociativos a partir de especificaciones de requisitos de software. Universidad de Salamanca. Proceedings of the II ADIS 2001 Workshop on Decision Support in Software Engineering. Almagro, Ciudad Real, Spain. ISBN 84-688-6649-0.

18. Moreno García M. N., Quintales L. A. ,
García Peñalvo F. J. y Polo Martín M.
J. (2002). Obtención y Validación de
Modelos de Estimación de Software
Mediante Técnicas de Minería de
Datos. (pp. 53-71). Revista
Colombiana de Computación. Volumen
3, nro 1.
19. Ramírez, E. L., Ambriz, M. H.,
Nungaray, M. A., Chessani, M. D. H.,
& Reyes, L. M. Á. R. (2012). Prototipo
para Almacenar y Recuperar Métricas
de Software. *Conciencia Tecnológica*,
43:11-17.

Ingeniería del Software

Requisitos de usuario y gestión de la demanda en AOP4ST

Fernando Pincioli

Instituto de Investigaciones
Facultad de Informática y Diseño
Universidad Champagnat
Mendoza, Argentina, +54 (261) 424-8443
pinciolifernando@uch.edu.ar

Resumen

El presente proyecto tiene como objetivo el establecer los lineamientos para la fase de requisitos de usuario del ciclo de vida del desarrollo de software perteneciente al proceso marco denominado Aspect-Oriented Process for a Smooth Transition (AOP4ST), de manera de que le permita:

- a) Permitir, ante todo, alcanzar los objetivos específicos de cada una de sus fases.
- b) Mantener un proceso homogéneo a lo largo de todas sus fases en cuanto a herramientas, estándares, notación, técnicas, etc.
- c) Colaborar con la detección de incumbencias en forma progresiva y manteniendo su separación a lo largo del ciclo de vida.
- d) Aportar a la definición de las mejores prácticas para la ingeniería de requisitos orientada a aspectos.

Este proyecto es el tercer proyecto bianual en cadena desarrollado en el Instituto de Investigaciones de la Facultad de Informática y Diseño de la Universidad Champagnat, dentro de la línea de desarrollo de software orientado a aspectos iniciada en el año 2013.

Palabras clave: orientación a aspectos, procesos de desarrollo de software, early aspects, modelado de negocio, modelado de requisitos de usuario, requisitos funcionales, atributos de calidad, reglas de negocio, separación de incumbencias, composición de incumbencias, resolución de conflictos, AOP4ST.

Contexto

El Instituto de Investigaciones de la Facultad de Informática y Diseño “Prof. Carlos Olivera”, de la Universidad Champagnat, tiene una línea de investigación en el área de Ingeniería de Software, y más específicamente en el desarrollo de software orientado a aspectos (AOSD). En ella se han desarrollado diversos proyectos de investigación. El primero de ellos, “*Procesos de desarrollo de software de calidad basados en aspectos*”, se llevó a cabo con la UTN Facultad Regional Mendoza [1], proyecto que luego continuó en nuestra Universidad Champagnat con el proyecto “*Definición de criterios para la detección temprana de aspectos en el modelado de negocios y el desarrollo de los requisitos*” llevado a cabo desde 2013 a 2016 y presentado en WICC 2015 [2]. Continuó con el proyecto, “*Modelado de procesos de negocio*”

orientados a aspectos con BPMN”, de 2016 a febrero de 2018, que tuvo como alcance la primera fase del ciclo de vida del desarrollo de software, al plantearse como objetivo el definir un proceso de modelado orientado a aspectos y que fue presentado en WICC 2016 [3] y 2017 [4]. Este nuevo proyecto continúa con la fase de requisitos de usuario y se inició en febrero de este año. A lo largo de todos estos proyectos se realizaron numerosas publicaciones en congresos y revistas nacionales e internacionales. Este proyecto cuenta con la financiación de la Universidad Champagnat y recibe aportes de la empresa Aconcagua Software Factory S.A., una de las principales fábricas de software de la provincia de Mendoza, con oficinas también en Buenos Aires, Santiago de Chile, Madrid y Barcelona.

Introducción

La definición de un proceso de desarrollo exige abarcar numerosos frentes. Uno de ellos corresponde a la definición específica de cada una de sus fases del ciclo de vida del desarrollo de software. En el presente proyecto se procura establecer los lineamientos de la fase de requisitos de usuario, que en la industria se denomina “gestión de la demanda” [5]. Al mismo tiempo, también se procura que esta fase continúe en forma homogénea y natural de la fase anterior, correspondiente al modelo de procesos de negocio.

En cuanto a las exigencias del paradigma en particular que se está empleando, se espera que en esta definición se tengan en cuenta las actividades propias del enfoque orientado a aspectos, que son la detección de las incumbencias, su separación, su encapsulamiento, la posterior composición y la resolución de los posibles conflictos.

Además, como esta fase del ciclo de vida es continuación de una anterior, es imprescindible el aseguramiento de la trazabilidad con las incumbencias detectadas en el modelo anterior.

Finalmente, también se espera establecer buenas prácticas de manera de asegurar los objetivos de la fase del ciclo de vida en cuestión, la calidad del producto software en desarrollo y la obtención de los beneficios que son la razón del empleo del paradigma de orientación a aspectos. Estos beneficios apuntan al desarrollo de un producto de software final, como así también de los productos intermedios que permiten producirlo, más modular, mantenible, reusable, extensible, comprensible, etc. [6], al administrar en forma separada las incumbencias que están *desparramadas* y *enredadas* en cada uno de los niveles de abstracción a lo largo del ciclo de vida completo.

Estas incumbencias pueden corresponder a cuestiones que son específicas del dominio del problema y también cuestiones que no lo son, lo que corresponde a los enfoques simétrico y asimétrico de la orientación a aspectos [7] [8].

Desde ya que los principales aportes a los requisitos tempranos (*early aspects*) son fuentes obligadas de referencia y de las que nos nutrimos en este proyecto [9] [10] [11] [12] [13]. También, los estudios de mapeo sistemático de Parreira [14] y de Narender [15] nos orientan en la organización y selección de las principales fuentes.

Líneas de Investigación, Desarrollo e Innovación

Los proyectos de investigación que se vienen sucediendo en nuestra Facultad

presentan cuatro ejes de investigación, donde los tres primeros se ven integrados con el cuarto:

1. *Modelado de procesos de negocio orientados a aspectos.*
2. *Separación y composición de incumbencias con resolución de conflictos.*
3. *Especificación de requisitos y gestión de incumbencias con casos de uso e historias de usuario.*
4. *Procesos de desarrollo de software orientados a aspectos.*

Resultados y Objetivos

Existen resultados alcanzados por los proyectos anteriores que posibilitan los resultados de este proyecto.

Los resultados provenientes de los proyectos anteriores son la definición del marco del proceso AOP4ST [1] [2] y la descripción detallada de ese proceso para su primer modelo, el modelo de procesos de negocio [3] [4] [16] [17] [18].

A partir de ellos se elaboraron las definiciones para el segundo modelo, el modelo de requisitos de usuario, que constituye el foco del presente proyecto.

Siguiendo las actividades que correspondientes a los modelos orientados a aspectos mencionados en el punto anterior, se avanzó en los siguientes puntos:

Detección y separación de incumbencias: en el modelo de negocio se detectaron incumbencias, que deben trasladarse al modelo de requisitos de usuario y constituyen las primeras incumbencias de este modelo. Al seguir los procesos de negocio del modelo inicial, se deben obtener los requisitos de usuario que pueden ser de tipo funcional, no funcional (atributos de calidad) y reglas de negocio. Los requisitos funcionales son más sencillos de obtener y su gestión cae

dentro del enfoque simétrico de AOSD. Ellos se van ubicando dentro de las incumbencias heredadas del primer modelo, pero pueden tener trazabilidad “muchos a muchos” ellas, lo que podría exigir la definición de una nueva incumbencia transversal en este modelo. Algunas ideas al respecto fueron presentadas en [19].

Los atributos de calidad, o requisitos no funcionales, siguen el mismo criterio empleado para los requisitos funcionales, pero por su propia naturaleza tienen relaciones de contribución positivas y negativas que obligan a rever las incumbencias a las que son asignados. Las consecuencias de estas relaciones de contribuciones fueron presentadas en [20] y [21].

Con respecto a las reglas de negocio, se siguen también los lineamientos descriptos y que, para la detección de incumbencias originadas por los tres tipos de requisitos, se publicaron los avances en [22].

Composición de incumbencias: no hemos avanzado por el momento en este punto, más que nada porque entendemos que con la definición de la trazabilidad se estaría cubriendo, pero entendemos que esto requiere un estudio más profundo y que está pendiente en nuestra agenda de trabajo.

Resolución de conflictos: también está pendiente de avance, aunque en esta fase del ciclo de vida, los conflictos entre requisitos han sido largamente estudiados y, además, el estudio entre las relaciones de contribución positivas y negativas mencionadas entre atributos de calidad, en el que hemos hecho avances importantes, también aportan a los objetivos de esta actividad.

Por último, está pendiente de publicación un estudio de mapeo sistemático sobre herramientas, técnica, notaciones, etc. orientadas a aspectos existentes a lo largo de la vida del

paradigma, cuyo protocolo está disponible en ArXiv [23].

Los criterios adoptados para esta fase del ciclo de vida están siendo aplicados en diversos proyectos -y para ser honestos, con diferente suerte, ya que dependemos de los casos que se presentan y no de los que deseáramos tener- de Aconcagua Software Factory S.A., una de las principales fábricas de software de la provincia de Mendoza.

Formación de Recursos Humanos

El proyecto apunta a la formación del equipo de profesores del área de Ingeniería de Software y de los alumnos y egresados de la Universidad Champagnat.

AOP4ST es el tema central de la tesis del doctorando Fernando Pinciroli, en el Doctorado en Ciencias Informáticas de la Universidad Nacional de San Juan, bajo la dirección del Dr. Raymundo Forradellas, de la Universidad Nacional de Cuyo, y la codirección del Dr. José Luis Barros Justo, de la Universidad de Vigo, España.

También se está elaborando la tesis de Gustavo Albino, de la Maestría en Ingeniería de Software de la Universidad Nacional de San Luis, y dos tesinas de grado de alumnos de la carrera de Licenciatura en Sistemas de Información de nuestra Universidad.

Referencias

- [1] F. Pinciroli, “Consideraciones para un proceso de desarrollo de software de calidad orientado a aspectos”. En: EnIDI 2011, Sexto Encuentro de Investigadores y Docentes de Ingeniería, San Rafael, 2011.
- [2] F. Pinciroli, “AOP4ST – Aspect-Oriented Process for a Smooth Transition”. En: WICC 2015, XVII Workshop de Investigadores en Ciencias de la Computación, Salta, 2015.
- [3] F. Pinciroli y L. Zeligueta, “El modelo de negocio en AOP4ST”. En: WICC 2016, XVIII Workshop de Investigadores en Ciencias de la Computación, Concordia, 2016.
- [4] F. Pinciroli y L. Zeligueta, “Modelado de negocios orientado a aspectos con AOP4ST”. En: WICC 2017, XIX Workshop de Investigadores en Ciencias de la Computación, Buenos Aires, 2017.
- [5] I. Alonso, J. Carrillo Verdún y E. Tovar Caro. “Designing an IT Artifact: Framework for the Entire Process of IT Demand Management”. En: Á. Rocha, A. M. Correia, H. Adeli, L. P. Reis y M. Mendonça Teixeira, “Analysis in Intelligent Systems and Technologies”, vol. 444, Springer, 2016.
- [6] A. Mendhekar, G. Kiczales y J. Lamping, “RG: A Case-Study for Aspect-Oriented Programming”. En: Technical Report SPL97-009P9710044, Xerox PARC, February, 1997.
- [7] J. Bálík y V. Vranić, “Symmetric aspect-orientation: some practical consequences”. En: Proceedings of the 2012 Workshop Next Generation Modularity, pp. 7–11, 2012.
- [8] W. Harrison, H. Ossher y P. Tarr, “Asymmetrically vs. Symmetrically Organized Paradigms for Software Composition”. En: IBM Research Report RC22685, 2002.
- [9] N. Weston, R. Chitchyan y A. Rashid, “Formal semantic conflict detection in aspect-oriented requirements”. En: Requirement Engineering, vol. 14, no. 4, pp. 247–268, 2009.
- [10] A. Y. Mohamed, A. E. F. Hegazy y A. Dawood, “Aspect Oriented Requirements Engineering”. En: Comput. Inf. Sci., vol. 3, no. 4, pp. 135–154, 2010.
- [11] A. Rashid, P. Sawyer, A. Moreira y J. Araújo, “Early aspects: A model for aspect-oriented requirements engineering”. En: Proc. IEEE International Conference on Requirement Engineering, vol. 2002–January, pp. 199–202, 2002.
- [12] S. Clarke y E. Baniassad. “Aspect-oriented analysis and design. The Theme

- approach”. Addison-Wesley, Boston, 2005.
- [13] I. Jacobson y P. Ng. “Aspect-oriented software development with use cases”. Addison-Wesley, New Jersey, 2005.
- [14] P. A. Parreira Júnior y R. A. Delloso Penteadó, “Aspect-Oriented Requirements Engineering. A Systematic Mapping”. En: XVI International Conference on Enterprise Information Systems, vol. 1, pp. 83–95, Lisboa, 2014.
- [15] N. Singh y N. Singh Gill, “Aspect-Oriented Requirements Engineering for Advanced Separation of Concerns: A Review”. En: International Journal on Computer Sciences, vol. 8, no. 5, pp. 288–298, 2011.
- [16] F. Pinciroli, J. L. Barros Justo y R. Forradellas, “Aspect-Oriented Business Process Modeling Approaches: An assessment of AOP4ST”. En: 46 JAIIO, Jornadas Argentinas de Informática, Córdoba, 2017.
- [17] F. Pinciroli, “Aspect-oriented business process composition rules in AOP4ST”. En: 35th International Conference of the Chilean Computer Science Society (SCCC 2016), held in conjunction with the 42th Latin American Computing Conference (CLEI 2016)., pp. 1–6, 2016.
- [18] F. Pinciroli, “Considerações acerca da mineração de aspectos”. Em: Perspectivas em Ciências Tecnológicas, vol. 5, no. 5, pp. 83–101, 2016.
- [19] F. Pinciroli y J. L. Barros Justo, “Early aspects in Aspect-Oriented Process for a Smooth Transition”. Em: XXIII Congreso Argentino de Ciencias de la Computación, La Plata, 2017.
- [20] F. Pinciroli, “Improving software applications quality by considering the contribution relationship among quality attributes”. En: Procedia Computer Sciences, 3rd International Workshop on Computer Antifragility Antifragile Engineering (ANTIFRAGILE 2016), vol. 83, pp. 970–975, 2016.
- [21] F. Pinciroli, “An HCI quality attributes taxonomy for an impact analysis to interactive systems design and improvement”. En: XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016), San Luis, 2016.
- [22] F. Pinciroli, “Concern detection along the requirement development”. En: CICCASI 2017, Congreso Internacional de Ciencias de la Computación y Sistemas de Información, Mendoza, 2017.
- [23] F. Pinciroli, L. Zeligueta, M. I. Lund y J. L. Barros Justo, “Systematic Mapping Protocol - Coverage of Aspect-oriented Methodologies for the Early Phases of the Software Development Life Cycle”. Technical Report, arXiv:1702.02653 [cs.SE], 2017.

Un Modelo de Análisis para aplicación de patrones de Buenas Prácticas en el Modelado Conceptual de Gobierno Electrónico

Oscar Carlos Medina, Marcelo Martín Marciszack, Mario Alberto Groppo
CIDS, Centro de Investigación, Desarrollo y Transferencia de Sistemas de
Información

Universidad Tecnológica Nacional, Facultad Regional Córdoba
omedina@frc.utn.edu.ar, marciszack@frc.utn.edu.ar, sistemas@groppo.com.ar

RESUMEN

El presente trabajo describe una propuesta de análisis que consiste en caracterizar patrones de Buenas Prácticas de E-Gobierno que se puedan utilizar en el Modelo Conceptual de los sitios web del sector público para que facilite su construcción con mayor nivel de calidad. La dimensión que se intenta optimizar en el presente trabajo es el nivel de calidad de sistemas de Gobierno Electrónico de Argentina, al no tener actualmente un modelo de solución exitoso para tomar como punto de partida. Al igual que en otros sistemas de información que resuelven problemáticas bien definidas, como por ejemplo los de ERP (acrónimo en inglés de Planeamiento de Recursos Empresariales), se identificó la necesidad de un conjunto de mejores prácticas para el Modelado Conceptual de los portales de ciudadanía digital, a las cuales pudiera referenciarse y valerse de ellas con las ventajas que ofrece el modelado con patrones en la construcción de un nuevo software público.

Palabras clave: Patrones, Buenas Prácticas, Modelo Conceptual, Gobierno Electrónico, E-Gob.

CONTEXTO

La preocupación por definir los requisitos de un sistema de manera

adecuada es extensamente tratada en [Sommerville 2005], donde el eje central es la definición de Buenas Prácticas en el establecimiento de los mismos, ya que plantea que *"el éxito de cualquier proyecto de desarrollo está íntimamente relacionado con la calidad de los requisitos"* y que *"el proceso de los requisitos es mucho menos homogénea y bien entendido que el proceso de desarrollo de software en su conjunto"*. El estudio de esta problemática desde el enfoque de la Ingeniería de Software continúa buscando la mejor forma de resolverlo como en [Wieggers & Beatty 2013] y [Pohl & Rupp 2015].

Para los sistemas de Gobierno Electrónico a principios de este siglo se consensuaron modelos de software público regionales. Como primeros pasos, nuestro país, veinte gobiernos latinoamericanos, España y el CLAD (Centro Latinoamericano de Administración para el Desarrollo), recomiendan en la "Carta Iberoamericana de Gobierno Electrónico": *"el uso de estándares abiertos y de software libre en razón de la seguridad, sostenibilidad a largo plazo y para prevenir que el conocimiento público no sea privatizado"* [CLAD 2007] y luego suscriben un "Modelo Iberoamericano de Software Público para el Gobierno Electrónico" en que se menciona como una de sus premisas las experiencias del gobierno federal brasileño *"tratar al software como*

un producto acabado que llega a la sociedad con documentación completa de instalación, y preparado para funcionar, como cualquier software" [CLAD 2010].

1. INTRODUCCIÓN

Siguiendo estas líneas de razonamiento, los autores del presente trabajo consideran que el Modelado Conceptual de los sistemas de Gobierno Electrónico en nuestro país pueden iniciarse en la implementación de mejores prácticas de un modelo regional de software público, para el que se propone utilizar patrones en la definición de las mencionadas mejores prácticas, con el valor agregado de su elaboración y despliegue. Los patrones de Buenas Prácticas no tienen solamente el alcance de una plantilla de sitio web, o un listado de sugerencias de diseño, sino que son soluciones de análisis y diseño concretas que sirven como guía en la etapa de construcción inicial de todo sistema con un nivel de calidad probado. En Argentina hay algunos proyectos que promueven el reconocimiento de las Buenas Prácticas de gobernanza como el del Banco de Experiencias Locales (Universidad Nacional de Quilmes y Universidad Nacional de General Sarmiento) pero aún no han cristalizado trabajos específicos en relación a las Buenas Prácticas de Gobierno Electrónico desde el presente enfoque. En cambio Brasil, en el 2008 presentó sus Estándares de Interoperabilidad de Gobierno Electrónico "e-PING" definiendo *"un conjunto mínimo de premisas, políticas y especificaciones técnicas que reglamentan la utilización de la Tecnología de Información y Comunicación (TIC) en la interoperabilidad de Servicios de Gobierno Electrónico, estableciendo las*

condiciones de interacción con los demás Poderes y esferas de gobierno y con la sociedad en general" [Gobierno Brasileño 2008]. Aunque estos estándares son principalmente técnicos, tienen el valor relevante de incorporar el concepto de patrones en la especificación de estándares. El desarrollo de sistemas de Gobierno Electrónico es considerado un problema de interés público en los mencionados acuerdos iberoamericanos gestionados por el CLAD, al que también suscribió y es partícipe activo Argentina, en concordancia con el nuevo paradigma de "gobierno abierto" que pregona Oscar Oszlak desde la red GEALC (Gobierno Electrónico de América Latina y el Caribe), quien advierte sobre el riesgo de que *"los sistemas de información suelen ser el talón de Aquiles de la responsabilización"* [Oszlak 2013] y que estos mecanismos de responsabilización ciudadana podrían ser candidatos también de mejores prácticas. Además, en Argentina el Gobierno Electrónico se incluye desde la última década dentro de un área de real interés científico y tecnológico [SECYT 2006].

Se considera que emplear el concepto de patrones en recomendaciones de prácticas como soluciones exitosas de procesos de negocio gubernamentales, es totalmente factible, por ejemplo trabajando con distintos tipos de patrones de diseño que son *"aplicables a un problema específico, dentro de un contexto determinado y que ofrezcan una solución clara al analista sobre el problema planteado"*. Con ellos *"se podría generar una solución que pueda adaptarse al nuevo entorno, reutilizando una solución ya probada"* [Appleton 2000]. La propuesta se desprende de un proyecto general, en desarrollo actualmente en el Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información (CIDS) de la

Universidad Tecnológica Nacional, Facultad Regional Córdoba, denominado "Implementación de Patrones en la validación de Modelos Conceptuales" (UTN-3604). Éste es continuación del proyecto anterior "Validación de Requerimientos a través de Modelos Conceptuales" (UTN-1643) donde se plantea incorporar en la actividad de Modelado Conceptual, esquemas predeterminados de análisis y diseño, para optimizar la definición inicial de un nuevo sistema, formalizando dicha definición, facilitando la validación de los requerimientos funcionales y reutilizando el conocimiento y experiencia de sistemas anteriores encapsulado en patrones de análisis y diseño.

Finalmente, para medir si se ha optimizado el nivel de calidad del software aplicando patrones de Buenas Prácticas de E-Gobierno, se va a recurrir a modelos de evaluación de calidad específicos para sistemas de Gobierno Electrónico desde distintas perspectivas acompañando la evolución de convertir procesos tradicionales a servicios en línea. Una reciente revisión bibliográfica sobre la evaluación de la calidad de los servicios de E-Gobierno [Sá, Rocha, Pérez Cota, 2015] propone un nuevo modelo de medición aplicable a los servicios de gobiernos locales conformado por dos niveles: dominios y dimensiones. Este trabajo de sistematización del conocimiento actual sobre el tema sintetiza en un esquema las dimensiones de evaluación de servicios de E-Gobierno, definiendo cada uno de ellos y citando las referencias científicas de estudio. Se consiga a continuación las dimensiones agrupadas por dominio sobre las que se evaluará cada sitio web público partícipe del proyecto:

a) Dominio Técnico: Usabilidad, Diseño y Calidad técnica.

b) Dominio Organizacional: Atención al cliente, Canales alternativos, Transparencia, Reclamos, Personalización, Rol político y E-Gobierno.

c) Dominio Seguridad: Privacidad, Seguridad, Confiabilidad y "Delivery" del servicio.

d) Dominio Información: Calidad de la información y Acción de informar.

Este modelo de descomposición del servicio de E-Gobierno en dominios y dimensiones puede tomarse como estructura conceptual de la definición de una Buena Práctica, contando con el beneficio adicional que luego de su utilización en el Modelado Conceptual, el calificador de cada dimensión serviría de indicador comparativo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Basado en el marco conceptual explicado en la Introducción, el objetivo de este trabajo es indagar si el Modelado de los sistemas de E-Gobierno de Argentina parte de una línea base concreta, probada y aceptada como la más adecuada en un estándar denominado Buenas Prácticas dentro de un modelo iberoamericano, optimiza su nivel de calidad. Resumiendo, esta investigación busca dar respuesta a los siguientes interrogantes:

a) ¿Qué son Buenas Prácticas de Gobierno Electrónico?

b) Las Buenas Prácticas de E-Gobierno ¿tienen una metodología de elaboración universalmente instituida? ¿Cuáles son sus principales características?

c) ¿Se recomiendan Buenas Prácticas como punto de partida para iniciar el desarrollo de sistemas de Gobierno Electrónico de Argentina e Iberoamérica?

Para lo cual se definen los siguientes objetivos del proyecto:

Objetivo general: Definir patrones a partir de Buenas Prácticas que puedan incorporarse en el Modelo Conceptual de sistemas de Gobierno Electrónico de Argentina, en el marco de un modelo iberoamericano de software público.

Objetivos específicos:

- Caracterizar el conjunto de Buenas Prácticas de Gobierno Electrónico más representativas en la actualidad y en un contexto iberoamericano.
- Describir las características básicas de los patrones de análisis y de diseño.
- Identificar los patrones de análisis y de diseño existentes que permitan definir Buenas Prácticas de E-Gob.
- Proponer la aplicación de patrones de Buenas Prácticas en el Modelo Conceptual de sistemas de Gobierno Electrónico.
- Evaluar la aplicación de estos patrones de Buenas Prácticas en sistemas de Gobierno Electrónico de Argentina.

El diseño metodológico de este trabajo tiene un enfoque cualitativo con estudios exploratorios de Buenas Prácticas de Gobierno Electrónico y de patrones de análisis y diseño dentro del marco teórico de la Ingeniería de Software. Se realiza un análisis descriptivo de los patrones de Buenas Prácticas de E-Gob que puedan aplicarse en el Modelo Conceptual de software público de Argentina.

Para la consecución de cada uno de los objetivos específicos se llevará a cabo el siguiente conjunto de actividades:

- a) Analizar estado del arte sobre el tema.
- b) Caracterizar mejores prácticas de E-Gobierno.
- c) Describir de patrones de análisis y diseño.
- d) Identificar patrones de análisis y diseño para Buenas Prácticas de E-Gobierno.

e) Proponer la aplicación de estos patrones en el Modelado Conceptual de sistemas de E-Gobierno.

f) Analizar y divulgar los resultados de la investigación.

3. RESULTADOS OBTENIDOS/ESPERADOS

Esta investigación proveerá al mercado informático no solo de mejores prácticas en lo que se refiere fundamentalmente a patrones de E-Gobierno, sino también de una metodología para asegurar la calidad del software desde etapas tempranas de su diseño y construcción a través del uso de patrones. A su vez garantizará la utilización de Buenas Prácticas y cierto grado de performance por la utilización de soluciones que ayudarían a tener en cuenta posibles requisitos necesarios para etapas posteriores del ciclo de desarrollo del software.

Las consultoras de software y las Universidades del ámbito tecnológico tendrán la posibilidad de acceder al fruto de este proyecto, lo que coadyuvará a mejores servicios en las diversas instituciones y organizaciones de la región al contar con patrones y Buenas Prácticas contribuyendo así al desarrollo productivo del software en el país.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto prevé la formación de recursos humanos que integrarán el mismo. Tendrá radicada la siguiente tesis doctoral que contribuirá con la ejecución del proyecto: “Definición de patrones a partir de Buenas Prácticas para el desarrollo de sistemas de Gobierno Electrónico”, para optar al grado académico de Doctor en Ingeniería,

Mención en Sistemas a realizar por el Ing. Oscar Carlos Medina y con dirección del Prof. Dr. Mario Alberto Groppo, codirigida por el Prof. Dr. Marcelo Martín Marciszack.

Se incorporará a becarios alumnos de investigación y a un becario Graduado BINID, los que tendrán a cargo la recolección y manipulación de datos y colaborarán en el desarrollo general del proyecto.

Al mismo tiempo, y como contribución a la formación de los integrantes, se elaborarán informes técnicos, como así también artículos para ser publicados en Congresos, Conferencias y reuniones científicas tanto a nivel nacional como a nivel internacional.

El equipo además de los docentes investigadores, estará conformado por estudiantes de la Carrera de Ingeniería en Sistemas de Información, con la finalidad de que inicien su formación en investigación científica y tecnológica, profundizando sus conocimientos en temas significativos de la especialidad. Los estudiantes tendrán la posibilidad de hacer la Práctica Supervisada, de quinto año, en el marco del proyecto. También se incorporará a docentes investigadores no categorizados que comenzarán a capacitarse en los procesos de investigación, y a graduados de la Carrera de Ingeniería en Sistemas de Información, con la finalidad de que inicien su formación en investigación científica y tecnológica, y pueda aportar a la comunidad los conocimientos adquiridos.

Iberoamericana de Gobierno Electrónico, IX Conferencia Iberoamericana de Ministros de Administración Pública y Reforma del Estado

CLAD Centro Latinoamericano de Administración para el Desarrollo (2010). Modelo Iberoamericano de Software Público para el Gobierno Electrónico”, Documento para la consideración de la XII Conferencia Iberoamericana de Ministros de Administración Pública y Reforma del Estado.

Gobierno Brasileño (2008). “e-PING. Estándares de Interoperabilidad de Gobierno Electrónico, Comité Ejecutivo de Gobierno Electrónico.

Oszlak, O. (2013). Gobierno abierto: hacia un nuevo paradigma de gestión pública, Red de Gobierno Electrónico de América Latina y el Caribe – Red GEALC.

Pohl, K. & Rupp, C. (2015). Requirements Engineering Fundamentals: A Study Guide for the Certified Professional for Requirements Engineering Exam, International Requirements Engineering Board.

SECYT Secretaría de Ciencia, Tecnología e Innovación productiva, Ministerio de Educación, Ciencia y Tecnología de la Nación Argentina (2006). Plan Estratégico Nacional de Ciencia, Tecnología e Innovación ‘Bicentenario’

Sá, F.A., Rocha, A., Pérez Cota, M. (2015). From the quality of traditional services to the quality of local e-Government online services: A literature review, Government information Quarterly N° 33, 149-160.

Sommerville, I., Sawyer, P. (2005). Requirements Engineering: A Good Practice Guide, Computing Department, Lancaster University, John Willey & Sons Ltd.

Wieggers, K. & Beatty, J. (2013). Software Requirements (Developer Best Practices), Microsoft Corp.

5. BIBLIOGRAFÍA

Appleton, B. (2000). Patterns and Software: Essential Concepts and Terminology.

CLAD Centro Latinoamericano de Administración para el Desarrollo (2007). Carta

Evaluación Temprana de la Usabilidad Empleando Patrones Embebidos en la Construcción del Modelo Conceptual para Aplicaciones Web.

Juan Carlos Moreno¹, Marcelo Martín Marciszack¹, María Alejandra Paz Menvielle¹, Claudia Castro¹, Andrea Delgado¹, Silvio Serra¹, Juan Pablo Fernández Taurant¹, Ezequiel E. Fernández¹, Gonzalo A. Saad¹, Nicolás Perez¹

Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información (CIDS)
Universidad Tecnológica Nacional - Facultad Regional Córdoba

Maestro López esq. Cruz Roja Argentina s/n, Ciudad Universitaria

{jmoreno33, marciszack, pazmalejandra, ingclaudiacastro, andreafdelgado, ingsilvioserra, jtaurant, ezequiel004, saad.gonzalo.ale, nicolasperezindorado}@gmail.com¹

RESUMEN

En el presente trabajo se propone llevar a cabo un estudio para el desarrollo de una metodología, que permita capturar aspectos de usabilidad en forma temprana en entornos de aplicaciones web. Se toma como base de partida los Modelos de Procesos de Negocios, los cuales mediante transformaciones ayudarán a obtener los modelos conceptuales correspondientes. En los modelos se detectarán patrones, que contribuirán a identificar e incorporar aspectos de usabilidad en etapas tempranas del ciclo de vida del software. A partir del modelo obtenido, se creará un conjunto de modelos abstractos que se utilizarán para validar los modelos conceptuales diseñados. Además, se obtendrá una trazabilidad del esquema conceptual diseñado, ante posibles errores, modificación o cambio de las especificaciones correspondientes, antes de que el modelo se encuentre terminado y se inicie la codificación del mismo.

Palabras clave: Modelos de Negocios, Escenarios, Patrones, MDA, Usabilidad.

CONTEXTO

La presente línea de Investigación constituye el fundamento de una tesis doctoral denominada “Metodología de evaluación temprana de la Usabilidad empleando patrones en la construcción del modelo conceptual de aplicaciones web”, la cual se encuentra inserta como proyecto colaborativo dentro del Proyecto de Investigación y Desarrollo “Un Modelo de Análisis para Aplicación de Patrones en el Modelado Conceptual de Aplicaciones Web” en el “Centro de Investigación, Desarrollo y Transferencia de Sistemas de

Información – CIDS” (Res. C.S.U. 2507/2016), bajo la Dirección del Prof. Dr. Ing. Marcelo Martín Marciszack y la Codirección de la Prof. Mg. Ing. María Alejandra Paz Menvielle, dependiente del Departamento Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Córdoba (código IAN4701 - Disp. SCTyP N° 252/2017). Para llevar a cabo este trabajo se empleará un método de investigación mixta o de triangulación, donde las unidades de observación estarán conformadas por expertos en desarrollo, y estudiantes universitarios en proceso de formación académica especializados en software.

1. INTRODUCCIÓN

El desarrollo de sistemas de información web se ha transformado en un proceso que busca construir aplicaciones útiles y correctas para su uso. Uno de los objetivos de la Ingeniería de Software es construir aplicaciones de calidad, útiles a los usuarios finales, aplicando distintos métodos y principios [1]. La calidad de las aplicaciones web se mide muchas veces basándose en el sentido común de los desarrolladores [2].

Por lo general, en el proceso de construcción del software se hace énfasis en los aspectos de la arquitectura, la funcionalidad y la persistencia de cada proceso, no tratándose de forma adecuada la interacción y facilidad de uso.

Por este motivo, el estudio la usabilidad del software web ha tomado relevancia. El concepto de usabilidad ha sido definido por varias normas de Organizaciones Internacionales de Estándares de Calidad (ISO, IEEE). En cada norma a la usabilidad está relacionado a la calidad del mismo.

La norma ISO/IEC 9126-1 [3], se considera a la usabilidad como un parámetro de calidad del software. Se reconoce a la usabilidad como “la capacidad en que un producto de software puede ser entendido, aprendido y usado por determinados usuarios bajo ciertas condiciones en un contexto de uso específico”. Se contempla la calidad interna, externa y en uso de un producto de software [4]. A su vez, la usabilidad es descompuesta en subatributos, haciendo que algunos atributos sean más tangibles y se puedan medir[5].

La norma ISO 25000 (Square) [6] contempla a la usabilidad bajo dos puntos de vista distintos: uno que contempla a la usabilidad desde el punto de vista del software, como producto en sí mismo; y el otro punto de vista desde la usabilidad de uso, desde la perspectiva del usuario.

A través de los distintos estándares se definen distintos atributos de la usabilidad, que sirven para formular métricas para la evaluación del software.

Pero la usabilidad es considerada en etapas finales de la construcción del software, cuando cualquier modificación afecta la arquitectura del sistema y el costo de cualquier modificación es alto [7], [8]. Una de las soluciones posibles a este problema, es incluir el análisis de la usabilidad en etapas tempranas, durante la fase de elicitación de los requisitos. Por esta razón se estudia el Entorno de Desarrollo de Software Dirigido por Modelos (DSDM) [9], también denominado MDD en el campo de la Ingeniería de Software, puesto que se busca saber si se considera la elicitación de requisitos de usabilidad en etapas de desarrollo tempranas de la construcción del software. En DSDM se busca la construcción de un software a través de una serie de modelos conceptuales que son independientes de la plataforma de implementación y representan del sistema de información. A través de estos modelos se busca generar el código final del programa, aplicando una serie de transformaciones.

La incorporación del Desarrollo de Software Dirigido por Modelos (MDD) al proceso de desarrollo de software, fue realizada por Object Management Group (OMG), fue denominada como Model-Driven Architecture (MDA) (Miller y Mukerji 2003) [10]. Para ello utilizó otros estándares, como el Lenguaje de Modelado Unificado (UML) para modelado, Facilidad de Meta Objetos (MOF para los metamodelos) y el lenguaje de intercambio de datos XML.

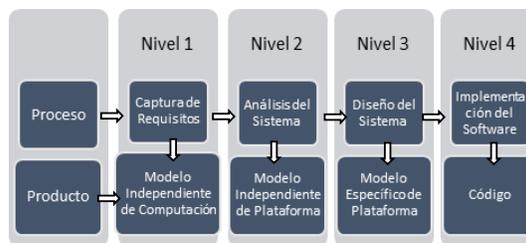


Figura 1. Model-Driven Architecture

Existen varios métodos de desarrollo de software de la Ingeniería web, que dan soporte al estándar MDD. Se pueden citar, como ejemplo, a los siguientes: OOHDM [11] [12], UWE [13] [14], OO-Method [15], OOH[16], OOWS[17], WebML[18].

El desarrollo de los sistemas Web en estos métodos, se lleva a cabo mediante modelos que capturan distintas vistas del sistema: un modelo estructural (modela contenido y comportamiento), un modelo de navegación (modela acceso al contenido) y un modelo de presentación abstracto (modela cómo el contenido es mostrado). La evaluación de estos modelos puede proporcionar información, para evaluar la usabilidad en etapas tempranas del desarrollo antes de que cualquier línea de código sea generada. Se debe considerar el nivel de abstracción de los modelos para poder evaluar las características de usabilidad.

Una vez obtenido un modelo conceptual (Casos de Usos, Diagramas de Interacción, etc), partiendo de un modelo de negocios sería interesante poder elaborar una serie de modelos abstractos donde se pueda aplicar la evaluación de Usabilidad y de accesibilidad simultáneamente. Esto es posible aplicando teoría de grafos empleando Autómatas Finitos Deterministas.

Pero también, el modelado debe proveer ciertos mecanismos que garanticen la incorporación de la usabilidad a través de buenas prácticas y soluciones ya probadas en la industria. Este razonamiento da origen a la utilización de patrones con aspectos de usabilidad incorporados que se emplearán para la construcción de los modelos conceptuales, que darán origen al software. Los patrones le otorgan eficiencia y reusabilidad a los modelos, por haber sido empleados para resolver problemas anteriores similares. Esto permite generar un catálogo de elementos reusables para el diseño del software.

Como consecuencia del Análisis del Modelo Conceptual y de haber seleccionado distintos artefactos del sistema de modelado, se debería analizar si es posible emplear patrones que incorporen la usabilidad en forma temprana, tanto para las interfaces como para el diseño de las bases de datos, para formular una nueva propuesta metodológica, que permita resolver la falta de usabilidad temprana.

Resultados y Objetivo

El objetivo general del presente trabajo será proponer una metodológico para la construcción de modelos conceptuales, que incorpore y evalúe la usabilidad en forma temprana a través de procesos estandarizados, empleando patrones para la incorporación de la usabilidad, y modelos abstractos para validar la presencia de la misma.

En función de lo anterior, se plantearon las siguientes metas:

- Describir los atributos y subatributos de usabilidad que debe poseer cualquier aplicación en entornos web. Describir requisitos de usabilidad básicos, que debe poseer cualquier aplicación web.
- Describir métricas de usabilidad, según la norma estándar de trabajo que se seleccione, que permitan medir y evaluar los resultados obtenidos.
- Comparar metodologías de desarrollo en entornos web, que incorporen Requerimientos Funcionales y No Funcionales en etapas tempranas del ciclo de vida.
- Caracterizar una nueva propuesta metodológica, describiendo los pasos necesarios para alcanzar el objetivo y subsanar las debilidades detectadas en otras metodologías.
- Identificar herramientas manuales o automatizadas, que ayuden a la consecución del objetivo, y sirvan como elemento de prueba empírica.
- Comparar la nueva propuesta metodológica para detectar fortalezas y debilidades.
- Identificar fortalezas y debilidades de la nueva propuesta, estableciendo distintas dimensiones de análisis, a los fines de poder elaborar una conclusión general de la propuesta metodológica y establecer futuras líneas de investigación.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO.

Las Líneas de Investigación de Desarrollo involucradas son:

- a) Modelos de Negocios: se busca mediante el análisis del modelo de negocios la captura de los requerimientos de usabilidad en forma temprana.
- b) Aplicación del desarrollo de software dirigido por modelos, mediante el cual se llevan a cabo transformaciones para obtener un modelo conceptual inicial.
- c) Obtención de escenarios y diccionarios de datos, con criterios de usabilidad definidos.
- d) Obtención de modelos abstractos mediante transformaciones. Los modelos abstractos representan a las interfaces que son sometidas a evaluación utilizando grafos para determinar la presencia de usabilidad.
- e) Identificar los distintos tipos de patrones aplicables para la incorporación de usabilidad en forma temprana en los modelos.

3. RESULTADOS ESPERADOS.

Como resultado general del proyecto se espera:

Establecer un marco teórico-metodológico para la incorporación de aspectos de usabilidad en forma temprana empleando patrones, a través del Desarrollo Dirigido por Modelos en una Metodología de Desarrollo Web y emplear modelos abstractos para su validación.

Asimismo, entre los resultados particulares se espera:

- Establecer requisitos de usabilidad básicos, que debe poseer cualquier aplicación web.
- Describir métricas de usabilidad aplicables en metodologías de desarrollo web, que soportan el Desarrollo de Software Dirigido por Modelos.
- Verificar la existencia de herramientas o factibilidad de construcción de las mismas.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro del marco del proyecto de investigación actual, en el CIDS, se están desarrollándose dos trabajos de Tesis de Doctorado en Ingeniería en Sistemas dentro de la U.T.N.-F.R.C.: uno pertenece a Juan Carlos Moreno, bajo la dirección del Dr. Marcelo Martín Marciszack, titulado “Metodología de evaluación temprana de la usabilidad empleando patrones en la construcción del modelo conceptual de aplicaciones web”; y otro perteneciente al Ing. Oscar Carlos Medina bajo la dirección del Dr. Mario Groppo cuyo tema es “Definición de patrones a partir de buenas prácticas para el desarrollo de sistemas de Gobierno Electrónico”. Además, dentro del proyecto, se está llevando a cabo la realización de una Tesis Posgrado de Maestría en Ingeniería de Software de la Universidad Nacional de la Plata, del Ingeniero Silvio Serra denominada “Modelos para la construcción de procesos de migración de datos en contexto de sistemas de desarrollo”, bajo la Dirección del Dr. Marcelo M. Marciszack y la Codirección del Dr. Gustavo Rossi. Por otra parte, la Esp. Ing. Claudia Castro se encuentra en proceso de análisis del material para la elaboración de su tesis de Maestría en Ingeniería en Sistemas de Información, en la Universidad Tecnológica Nacional de la Facultad Regional Córdoba. Sumado a esto, se han incorporado a docentes de la carrera Ing. en Sistemas de Información como investigadores de apoyo, a dos becarios Alumnos de Investigación y a un becario Graduado BINID que colaboran investigando en el desarrollo del proyecto.

Además, se conformó un proyecto integrador con otras Regionales (Villa María, Santa Fe, Resistencia y Córdoba). El proyecto está dentro de un PID INTEGRADOR PID IPN-4409, junto con otros 7 proyectos afines para realizar actividades en conjunto denominado “Herramientas y Métodos de soporte a la Ingeniería de Software: requerimientos, estrategias ágiles y calidad de procesos y productos” cuyo Director es el Prof. Dr. Horacio Pascual León.

Y finalmente se firmó un convenio de colaboración recíproca entre el Grupo de Investigación SII-GEAC de la Universidad de Vigo, bajo la Dirección del Prof. Dr. Manuel Pérez Cota de la Universidad de Vigo, en donde el

director de este proyecto forma parte integrante del mismo, y del CIDS de la Universidad Tecnológica Nacional - Facultad Regional Córdoba, en donde se desarrolla este proyecto de investigación. Por medio de este convenio los integrantes de ambas Instituciones se comprometen a colaborar en forma mutua.

Los resultados y conclusiones arribados en estos proyectos se publican y exponen mediante artículos en Congresos, Conferencias y revistas científicas tanto a nivel nacional como a nivel internacional.

5. BIBLIOGRAFÍA

- [1] Pressman R., "What a tangled Web we weave," *IEEE Software*, 2000.
- [2] Abrahao S., Condori-Fernandez N., Olsina L., and Pastor O., "Defining and validating metrics for navigational models," Australia, 2003.
- [3] Norma ISO/IEC ISO9126-1, "Software Engineering -Product Quality - Part 1," 2001.
- [4] Nigel Bevan, "Quality and usability: A new framework," *Achieving software product quality*, 1997.
- [5] Mario G. Piattini, Felix O. Garcia, and Ismael Caballero, "Calidad de Sistemas Informáticos," México, ISBN 978-970-15-1267-8, 2007.
- [6] ISO/IEC 25000, Software Engineering - Software Product Quality Requirements and Evaluation (SQuaRE).
- [7] L Bass and B John, "Linking usability to software architecture patterns through general scenarios," *The journal of systems and software*, no. 66, pp. 187 - 197, 2003.
- [8] Eelke Folmer and Jan Bosh, "Architecting for usability: A survey. ," *Journal of Systems and Software*, pp. 61 - 78, 2004.
- [9] Stephen J Mellor, Kendall Scott, Axel Uhl, and Dirk Weise, *Model-Driven Architecture*. Berlin / Heidelberg: Springer, 2002.
- [10] (2003) MDA_Guide_Version1-0. [Online]. http://www.omg.org/mda/mda_files/MDA_Guide_Version1-0.pdf
- [11] G. Rossi and D. Schwabe, "Modeling and Implementing Web Applications using OOHDM," in *Web Engineering, Modeling and Implementing*

- Web Applications.*: Springer, 2008, pp. 109-155.
- [12] S. Daniel, P. Rita de Almeida, and M. Isbela, "OOHDM-Web: an environment for implementation of hypermedia applications in the WWW," in *SIGWEB News L8, 2*, 1999, pp. 18-34.
- [13] Nora Koch and Martin Wirsing, *Software Engineering for Adaptive Hypermedia Applications*. München, Germany: Ludwig-Maximilians University of Munich, 2000.
- [14] N. Koch, A. Knapp, G. Zhang, and H. Baumeister, "UML-Based Web Engineering, An Approach Based On Standar.," in *Web Engineering, Modelling and Implementing Web Applications.*: Springer, 2008, pp. 157-191.
- [15] Oscar Pastor and Juan Carlos Molina, *Model-Driven Architecture in Practice: A Software Production Environment Based on Conceptual Modeling*, Inc. Secaucus, NJ, Springer-Verlag New York, Ed. Valencia, USA, 2007.
- [16] J. Gómez and C. Cachero, "OO-H Method: extending UML to model web interfaces.," in *Information Modeling For internet Applications.* Hershey, PA.: Ed. IGI Publishing, , 2003, pp. 144-173.
- [17] P.V., Albert M., and Pastor O. Fons J.,: LNCS. Springer, 2003, vol. 2813, pp. 232-245.
- [18] S. Ceri, P Fraternali, and A. Bongio, "Web Modeling Language (WebML): a modeling language for designing Web sites.," in *9th. World Wide Web Conference*, 2000, pp. 137-157.

ANÁLISIS DE FRAMEWORKS WEB ADAPTATIVOS BASADOS EN HTML5

Rocío A Rodríguez, Pablo M Vera, Isabel B Marko, Gaspar Acevedo Zain, Claudia G Alderete

GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software)
Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Florencio Varela 1903, San Justo, Provincia de Buenos Aires
{rocio.rodriguez, pvera, imarko, gacevedo, calderete} @unlam.edu.ar

RESUMEN

Actualmente se dispone de una gran cantidad de dispositivos móviles con posibilidades de navegar por la web. Mayormente los teléfonos celulares son el dispositivo móvil más utilizado para este fin. Tomando en consideración los smartphones, tablet y otros dispositivos es necesario planificar una web que pueda adaptarse a las diversas pantallas evitando que el usuario tenga que hacer zoom y scroll en ambas direcciones para poder visualizar los contenidos. En este sentido el diseño adaptativo resuelve el problema de visualización. Sin embargo, no todos los frameworks basados en HTML 5 tienen las mismas posibilidades, es por ello que resulta de interés analizar las características de los frameworks existentes.

En el presente proyecto se pone el foco en la calidad de los sitios web adaptativos creados con frameworks basados en HTML 5. Para esto es necesario conocer el funcionamiento de dichos frameworks, con la finalidad de poder construir métricas que permitan luego evaluar la calidad de la solución final.

Palabras clave: Frameworks Web Adaptativo, Dispositivos Móviles, Usabilidad

CONTEXTO

El presente proyecto de Investigación y Desarrollo corresponde al grupo de investigación GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software) perteneciente al Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de

La Matanza (UNLaM). El grupo es interdisciplinar al estar formado por docentes de distintos departamentos y en él se encuentran formándose en actividades de I+D alumnos de la universidad de las carreras de Ingeniería.

1. INTRODUCCIÓN

La alta inserción de dispositivos móviles y el acceso desde los mismos a la web es un tema que debe ser considerado al momento de diseñar una solución web. “Actualmente existe una clara tendencia en el crecimiento del uso de los dispositivos móviles para conectarse a internet ... Sin embargo, al querer ingresar a muchos sitios web desde los dispositivos móviles surgen diversos problemas ya que la mayoría de los sitios no fueron diseñados y estructurados pensando en dispositivos móviles sino para ser visualizados desde una computadora” [1]. El dispositivo móvil más frecuentemente utilizado para acceder a la web son los teléfonos celulares. “En Argentina la cantidad de líneas activas de telefonía móvil, supera al total de habitantes, habiendo una inserción del 122% (valor calculado tomando en cuenta la cantidad de habitantes según el Instituto Nacional de Estadísticas y Censos (INDEC) [2] y la cantidad de líneas activas informadas por la Comisión Nacional de Comunicaciones (CNC) [3])” [4].

Es por esto que resulta imprescindible tomar en cuenta a los dispositivos móviles y planificar un diseño centrado en los mismos, lo cual puede efectuarse realizando una

solución particular para ellos o bien una solución adaptativa.

“El diseño web adaptativo utiliza hojas de estilo en cascada complejas, para modificar de forma fluida la apariencia de un sitio web dependiendo del ancho de la ventana de visualización del dispositivo en uso” [5].

“El diseño adaptativo se basa en la utilización de hojas de estilo de nivel 3 con CSS Media Queries [6] lo que permite aplicar distintos estilos según sea el tamaño de la pantalla en la cual se está visualizando el sitio web. De esta forma, una misma página se podrá ver “correctamente” indistintamente del tamaño en la que se esté visualizando” [1].

Para simplificar la tarea existen diversos frameworks basados en HTML 5 que permiten construir sitios web adaptativos.

En base a Google Trends [7], Bootstrap [8] y Foundation [9] son los más populares. “Entre los frameworks más populares para diseño web adaptativo podemos encontrar el conocido Bootstrap, popularizado por ser usado para el desarrollo de la red social Twitter. Foundation pasa por ser uno de los frameworks más potentes y es la base de sitios como Pixar o National Geographic. Skeleton, por su parte, está considerado como uno de los más ligeros y más fáciles de personalizar, pues carece, a diferencia de los anteriores, de una gran profundidad en la definición visual de estilos, ofreciendo la estructura básica de desarrollo”[10].

En la figura 1, la línea superior muestra la tendencia de búsquedas de Bootstrap, la línea siguiente corresponde a Foundation, luego puede observarse que los frameworks restantes tienen tendencias bastante cercanas entre sí pero alejadas a estos dos frameworks principales. En la figura 2 puede observarse como Bootstrap y Foundation, son los frameworks más populares, particularmente en la vista de 5 años puede observarse que el liderazgo ha cambiado antes Foundation tenía una presencia superior a Bootstrap, lo cual se

ha revertido. De todos modos no cabe duda que son los dos frameworks que sobrepasan ampliamente a los restantes.

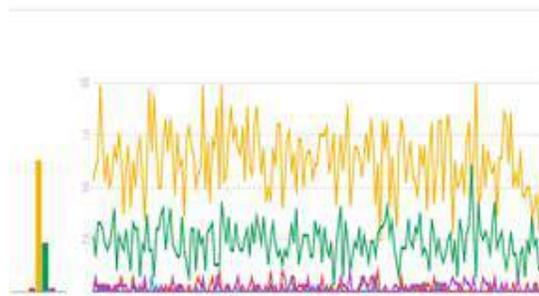


Figura 1. Frameworks más populares (tendencia actual)– Línea superior Bootstrap, línea siguiente Foundation, luego el resto de los frameworks

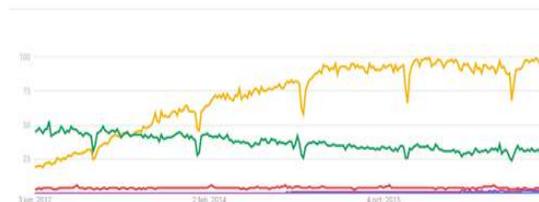


Figura 2. Frameworks más populares (vista últimos 5 años) – Inicialmente la tendencia era Foundation, lo cual ahora se ha revertido con Bootstrap

La tabla 1 presenta diversos frameworks ordenados por su tendencia de consulta, en base a google trends.

Tabla 1. Frameworks basados en HTML 5 ordenados por tendencia

<i>Tendencia</i>	<i>Framework</i>
1	Bootstrap [8]
2	Foundation [9]
3	Materialize CSS [11]
4	Skeleton [12]
5	Semantic ui [13]
6	Pure CSS [14]
7	Milligram [15]
8	w3css [16]

Al utilizar los frameworks mencionados en la tabla 1, pudo advertirse que la mayoría de ellos no están correctamente diseñados, ya que se basan en unidades absolutas (píxeles) para determinar los distintos tamaños de pantalla cuando lo correcto sería utilizar unidades relativas (em o porcentajes) para

asegurar una correcta visualización en pantallas pequeñas más allá de la resolución que posean. “Trabajar con proporciones en lugar de pixeles, en el posicionamiento de los componentes del sitio, marca un cambio sustantivo para su despliegue en áreas cambiantes o pantallas diversas” [17].

“Para que un sitio sea usable cuando se ve en una variedad de tamaños de pantalla, un cierto grado de flexibilidad debe ser incorporada en el diseño. Div, fuentes, botones, y todos los demás elementos deben ser capaces de crecer y contraerse sobre la base del tamaño de la pantalla. En el diseño web adaptativo, esto se logra mediante el abandono de los pixeles, en favor de porcentajes o unidades relativas llamadas ems” [18].

Es por ello que resulta interesante poder evaluar los sitios web adaptativos desde distintos aspectos que permitirán compararlos entre sí.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Se investigará principalmente sobre los siguientes temas:

- Evaluación de frameworks adaptativos
- Análisis de las soluciones realizadas con frameworks existentes
- Definición de principios y pautas de usabilidad
- Construcción de Métricas para evaluar sitios web adaptativos

3. RESULTADOS OBTENIDOS/ESPERADOS

Si bien en la muestra se consideraron frameworks livianos con menos potencia y recursos, todos los frameworks cuentan con controles que permiten mejorar la usabilidad en pantallas reducidas. Un aspecto importante y distintivo es la cantidad de tamaños que pueden definirse para visualizar los sitios en distintos dispositivos. Todos los framework adaptativos ofrecen diversos tamaños para los cuales pueden definirse la forma en que se

mostrarán los distintos elementos del sitio. El problema es que si bien la mayoría de los frameworks soportan estas vistas al momento del desarrollo no siempre se las considera ó si se las considera no se prueba verdaderamente como se visualizará la solución en los distintos tamaños de pantalla, encontrándose una gran cantidad de problemas (como por ejemplo: inconvenientes con el scroll, superposición de elementos, etc.), los cuales atentan contra la usabilidad. Esto se evalúa por observación e inspección del código fuente.

Actualmente los frameworks más utilizados tienen entre 4 ó 5 tamaños de dispositivos previstos (ver tabla 2). Estos tamaños se pueden implementar agregando instrucciones propias de los frameworks que permiten realizar puntos de corte o ruptura, los cuales determinarán el momento en que se cambiará la visualización. Es decir, tener 4 tamaños distintos previstos, permitirá tener 4 vistas distintas del mismo sitio web. Cuán eficaz sea cada una de dichas vistas y su impacto en la usabilidad dependerá de la implementación.

Tabla 2. Tamaños previstos por cada Frameworks

Framework	Tamaños Previstos
Bootstrap	4
Foundation	5
MaterializeCSS	4
Skeleton	5

Se implementó luego una misma solución con los distintos frameworks, la cual tuviese ciertas características básicas (imagen adaptativa, texto encolumnado, menú que se colapse con ícono en tamaño reducido...), con 3 vistas distintas, adaptándose a cada una de ellas. Dado que se han seleccionado todos los frameworks mencionados previamente en la tabla 2 se opta por 3 vistas porque hay frameworks que no cuentan con 4 ó más vistas y el objetivo es realizar una solución equivalente en todos ellos. Para realizar dicha tarea se establece un modelo a seguir (ver figura 3). Se desarrolló la solución para la página principal, del sitio modelo, en los distintos frameworks lo que permite comparar

la forma de codificación y la facilidad para realizar este objetivo.

Si bien la mayor parte de los frameworks cuentan con buena documentación y ejemplos que permiten realizar la solución planteada sin grandes dificultades. Las principales deficiencias encontradas fueron principalmente la dificultad en algunos casos para que verdaderamente se vean bien los contenidos ya que debe realizarse un esfuerzo en las pruebas e ir agregando algunos tags al lenguaje de etiquetado que permitan por ejemplo que en tamaño pequeño siga viéndose centrado un título, etc. Es decir, no debe confiarse en que por simplemente usar un framework adaptativo con controles que permiten la adaptación en diversos tamaños la solución se visualizará correctamente. Por otra parte, se ha detectado que casi la mitad frameworks utiliza unidades absolutas en vez de unidades relativas (ver tabla 3), lo cual trae problemas serios (por ejemplo, cuando un usuario cambia el tamaño de letra por defecto, dado que no verá reflejado en la página dicho cambio).

Tabla 3. Tamaños previstos por cada Frameworks

Framework	¿Usa unidades relativas?
Boostrap	SI
Foundation	SI
Materialize CSS	NO
Skeleton	SI
Semantic ui	NO
Pure CSS	SI
Milligram	SI
w3css	NO

Si bien el W3C cuenta con lineamientos para asegurar una buena usabilidad en sitios web basados en HTML 5 [19], los frameworks no han implementado todas estas buenas prácticas y muchas veces esto hace como consecuencia una mala implementación por parte de los desarrolladores. En base a esto se ha realizado un relevamiento sobre los sitios web adaptativos de instituciones superiores de Argentina y se ha evidenciado diversos problemas al navegarlos.

Con lo realizado hasta el momento, puede concluirse que los frameworks no están correctamente diseñados ya que se basan en unidades absolutas (píxeles) para determinar los distintos tamaños de pantalla cuando lo correcto sería utilizar unidades relativas (em o porcentajes) para asegurar una correcta visualización en pantallas pequeñas más allá de la resolución que posean.

Un mal uso del diseño adaptativo incorporará diversos errores, entre ellos:

- Páginas muy pesadas;
- Gran cantidad de imágenes;
- Imágenes no preparadas para dispositivos móviles;
- Contenido no priorizado;
- Páginas demasiado extensas.

Por lo que resulta sumamente importante poder construir métricas que permitan evaluar y ofrecer reportes que permitan mejorar los sitios web adaptativos.



Figura 3. Modelo de sitio web principal realizado con 3 vistas

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por 7 personas.

- 4 Docentes pertenecientes a diferentes departamentos conformando un grupo interdisciplinario.
- 1 Graduado en Ingeniería informática
- 2 alumnos de la carrera Ingeniería Informática quienes se están formando en actividades de Investigación y Desarrollo los cuales cuentan con becas asignadas

Vinculado con esta temática se encuentra en realización una tesis de doctorado en la Universidad Nacional de La Plata.

5. BIBLIOGRAFIA

- [1] Rodríguez, R. A., Marko, I., Vera, P. M., Valles, G. Y., & Acevedo Zain, G. (2017, August). Evaluación de la usabilidad en sitios web móviles adaptativos. In XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires).
- [2] INDEC – Instituto Nacional de Estadísticas y Censos, Argentina. <http://www.indec.mecon.ar/webcenso/index.asp>
- [3] CNC Comisión Nacional de Comunicaciones. La Telefonía Móvil en Argentina <http://www.cnc.gov.ar/ciudadanos/telefoniamovil/index.asp>
- [4] Rodríguez, R. A., Giulianelli, D. A., Trigueros, A., Vera, P. M., & Marko, I. (2010). Los organismos gubernamentales y la estrategia de la utilización de mensajes de texto para ofrecer m-Servicios. In XVI Congreso Argentino de Ciencias de la Computación.
- [5] Walsh, T. A., McMin, P., & Kapfhammer, G. M. “Automatic Detection of Potential Layout Faults Following Changes to Responsive Web Pages”. In Automated Software Engineering (ASE), 2015 30th IEEE/ACM International Conference on (pp. 709-714). IEEE.
- [6] W3C Media Queries. 2012. <https://www.w3.org/TR/css3-mediaqueries/>
- [7] Google Trends. <https://trends.google.com.ar/trends/>
- [8] Otto M., Thornton J. “Bootstrap”. 2010. <http://getbootstrap.com/>
- [9] Zurb. “Foundation - The most advanced responsive front-end framework in the world”. 2011. <http://foundation.zurb.com/>
- [10] Prego Ivan. Compara los Mejores Frameworks Adaptativos. 2014. <http://ivanprego.com/desarrollo-web/bootstrap-foundation-skeleton-compara-los-mejores-frameworks-para-diseno-adaptativo-o-responsive-web-design/>
- [11] Wang A, Chang A, Mark A, Louie K, “Carnegie Mellon University” <http://materializecss.com/>
- [12] Gamache D. “Skeleton”. 2014. <http://getskeleton.com/>
- [13] Semantic ui “User Interface is the language of the Web”. <https://semantic-ui.com/>
- [14] Pure CSS. “A set of small, responsive CSS modules that you can use in every web project”. <https://purecss.io/>
- [15] Milligram. “A minimalist CSS frameworks”. V1.3.0. <http://milligram.io/>
- [16] w3css. “W3.CSS Tutorial”. <https://www.w3schools.com/w3css/default.asp>
- [17] Labrada Martínez, E., & Salgado Ceballos, C. (2013). Diseño web adaptativo o responsivo. Tema del mes.
- [18] Guerra, Y. M., González, R. C., Febles, J. P. “Diseño web adaptativo para la plataforma educativa ZERA”. Revista Cubana de Ciencias Informáticas, 10(2), pp. 100-115. 2016
- [19] W3C. “Mobile Web Application Best Practices”. <https://www.w3.org/TR/mwabp/>

Bases de Datos de Proyectos de Software Ágiles para ser Utilizadas en Simulación de Proyectos.

Santiago Hernán Bareiro^a, Diego Alberto Godoy^b, Eduardo O. Sosa^c, Edgardo A. Belloni^d, Juan de Dios Benítez^e, Sergio Fidelis^f, Fabián Favret^g

Centro de Investigación en Tecnologías de la Información y Comunicaciones (C.I.T.I.C.)
Departamento de Ingeniería y Ciencias de la Producción-Universidad Gastón Dachary
Av. López y Planes 6519- Posadas, Misiones, Argentina. Teléfono: +54-376-4438677

^ahbareiro@citic.ugd.edu.ar, ^bdiegodoy@citic.ugd.edu.ar, ^ceduardo.sosa@citic.edu.ar,
^debelloni@ugd.edu.ar, ^ejuan.benitez@citic.ugd.edu.ar, ^fsfidelis@citic.ugd.edu.ar,
^gfabianfavret@citic.ugd.edu.ar

Resumen

En este trabajo se presenta un proyecto de investigación denominado “Simulación en las Tics: Diseño de Simuladores de Procesos de Desarrollo de Software Ágiles y Redes De Sensores Inalámbricos para la Industria y la Academia”, cuyo objetivo es diseñar simuladores de procesos de desarrollo de software ágiles y de redes de Sensores Inalámbricos para la Industria y la academia. Particularmente en este artículo se presentan los avances realizados en relación a construcción de un Sistema de Gestión de Datos Postmortem de Proyectos de Desarrollo de Software Que Utilizan Metodología Scrum.

Palabras claves: Simulación; Proyectos de Desarrollo de Software; Base de datos.

Contexto

El trabajo presentado en este artículo tiene como contexto marco el proyecto de investigación denominado “Simulación en las Tics: Diseño de Simuladores de Procesos de Desarrollo de Software Ágiles y Redes De Sensores Inalámbricos para la Industria y la Academia”, registrado actualmente en la Secretaría de Investigación y Desarrollo de la Universidad Gastón

Dachary (UGD) con el número Código IP A07003 y radicado en el Centro de Investigación en Tecnologías de la Información y Comunicaciones de dicha universidad.

El mismo fue incorporado como proyecto aprobado en el llamado a presentación interna de la UGD de proyectos de investigación N°7 mediante la Resolución Rectoral 07/A/17 y es una continuidad del Proyecto Simulación como herramienta para la mejora de los procesos de software desarrollados con metodologías ágiles utilizando dinámica de sistemas, R.R. UGD N° 18/A/14 y R.R. UGD N° 24/A/15.

Entre las líneas con mayores resultados dentro del proyecto referido, se encuentran las de: (i) “Modelo de Simulación Dinámico de Gestión de Proyectos de Desarrollo de Software que utilizan Programación Extrema”, (ii) “Modelo de Simulación Dinámico de Gestión de Proyectos de Desarrollo de Software Bajo Scrum” y (ii) “Modelo de Simulación Dinámico de Gestión de Proyectos de Desarrollo de Software desarrollados con Crystal”. Este artículo se enfoca en la presentación de un línea para estudiar la calendarización de proyectos llevados a cabo Utilizando XP [1] y diseñar un prototipo de herramienta de

calendarización de liberaciones e iteraciones en proyectos con XP.

1. Introducción

Generalmente, en el proceso de desarrollo de software se hace énfasis sobre el control de procesos mediante una rigurosa definición de roles, actividades y artefactos incluyendo modelado y documentación muy detallada. Como se explica en [2], este esquema tradicional para llevar a cabo el desarrollo de proyectos de software ha demostrado ser efectivo y necesario en proyectos de gran tamaño respecto a tiempo y recursos, donde en general, se exige un alto nivel de burocracia en el proceso. Sin embargo, este enfoque no resulta ser el más adecuado para proyectos actuales donde el contexto del sistema es muy cambiante y se exige reducir drásticamente los tiempos de desarrollo manteniendo una alta calidad en el producto. Como alternativa a los métodos tradicionales, con el objetivo de poder adaptarnos a ese ámbito cambiante, de mantenimiento de alta calidad del producto y reducción en los tiempos de desarrollo surgen las llamadas metodologías ágiles. Según un estudio realizado por la Scrum Alliance [3] [podría referencia al estudio acá], Scrum, es una de las metodologías ágiles más utilizadas actualmente, cerca del 95% de los encuestados utilizan prácticas de Scrum en la gestión de proyectos de software ágiles.

Scrum posee muchas características, como ser sus valores, principios y prácticas que hacen que la administración de proyectos de software usando esta metodología se torne compleja e impredecible debido a la gran cantidad de variables que el administrador del proyecto debe manejar. Ante estas circunstancias es posible utilizar modelos de simulación que permitan evaluar los resultados de diferentes decisiones de gestión sin

interferir en el desarrollo real del proyecto.

Los modelos de simulación de procesos de desarrollo de software permiten a los administradores de proyectos realizar cambios en variables críticas del proyecto como ser la cantidad de requerimientos a desarrollar en cada iteración, el tiempo de entrega de las mismas, la cantidad de programadores, las horas de trabajo por día, etc. De esta forma se puede evaluar el impacto de las decisiones de gestión, compararlas entre si y escoger la que mejor se ajusta al proyecto real.

Para poder aprovechar los beneficios de la simulación enunciados en el párrafo anterior con respecto a la gestión de proyectos de software mediante Scrum, hay que tener en cuenta que, todo modelo de simulación debe ser validado antes de poder ser aplicado en un experimento, como bien se explica en las etapas de la simulación propuestas en [4]. Para poder validar un modelo de simulación es necesario contar con información previa del sistema sobre el cual se requiere estudiar su comportamiento, con el fin de asegurar que dicho modelo es representativo del sistema que se modela. Como se menciona en [5], la falta de datos es un problema conocido en la simulación de la ingeniería de software, tanto [6] y [7] resaltan que resulta difícil encontrar información referida a datos de proyectos de software que hayan sido llevados a cabo mediante metodologías ágiles, por lo tanto, esto dificulta la validación de los modelos de simulación centrados en estudiar la gestión de este tipo de proyectos.

Dentro de la gestión de proyectos de desarrollo ágiles y sobre todo en Scrum, los datos más importantes, relacionados con la gestión del proyecto son [9] [10], los referidos a los costos, el esfuerzo, el tamaño del producto y la duración de

los sprints o iteraciones, los cuales son muy importantes porque están implicados en tareas de estimaciones. Donde la precisión de dichas tareas es fundamental para el éxito del proyecto, de lo contrario se podrían producir ciertos efectos negativos como ser [9] excesos de presupuestos, entregas que no se realizan a tiempo, mala calidad y baja aceptación del producto por parte del cliente.

Actualmente, existen herramientas que si bien son utilizadas para realizar estimaciones, no se centran específicamente en recopilar información histórica para validaciones de modelos de simulación basados en procesos de desarrollo de software ágiles. Solo permiten almacenar datos de proyectos ya desarrollados, con el fin de ir mejorando y precisando las estimaciones realizadas. Dentro de estas herramientas tenemos a Agile Cost Management Tool [11], CostEs [12] y Estimador de Proyectos Informáticos [13]. La desventaja de las tres herramientas mencionadas es que los datos de los proyectos deben ser ingresados manualmente por el usuario. Ninguna de estas herramientas permite que los datos necesarios para realizar las estimaciones, sean almacenarlos como datos históricos o puedan ser importados por otro software. Dentro de estas herramientas tenemos una gran variedad de software, entre las más utilizadas están, Atlassian [14] Jira, Ice Scrum [15] y Trello [16]. Las cuales permiten compartir información de los proyectos que se gestionen a través de éstas, mediante archivos de formatos libres para el intercambio de datos XML [14] (Atlassian Jira), CSV [15] (Ice Scrum) y JSON [17] (Trello). Esta característica de intercambio permite procesar los archivos de manera automática, extraer información relevante de los mismos y almacenarlos en una base de datos.

Con la existencia de una herramienta que permita extraer información de sistemas de gestión de proyectos que contribuya con generación de una base de datos histórica de proyectos anteriores ya desarrollados, sería mucho más sencillo validar los modelos de simulación que se enfoquen en estudiar el comportamiento de la gestión de proyectos de software ágiles que utilicen Scrum como metodología.

2. Línea de Investigación

Como objetivo general se propone: Diseñar una aplicación que interprete un tipo de formato para el intercambio de datos manejado por herramientas de gestión de proyectos de desarrollo de software basados en Scrum, con el fin de obtener información destinada a la validación de modelos de simulación de este tipo de proyectos.

Como objetivos específicos se realizarán los siguientes:

- Analizar cuáles son las variables de todo proyecto de desarrollo de software que utilice Scrum.
- Determinar cómo los modelos de simulación capturan los datos que se utilizan en el modelo.
- Estudiar tres de las herramientas más utilizadas para la gestión de proyectos Scrum.
- Diseñar la arquitectura de la aplicación.
- Validar el comportamiento de la aplicación conforme a cada una de las herramientas de gestión para las que se desarrolle.

3. Resultados

Esta sección tiene como propósito definir cuáles son las variables presentes en un proceso de desarrollo de software que utilice Scrum como metodología. Es importante conocer la identidad de estas variables para que se puedan tomar como referencia al

momento de determinar cuáles son las más utilizadas por los trabajos de simulación que se centren en el estudio de procesos Scrum y que, se vean reflejadas en el software propuesto en este trabajo. En base al análisis teórico realizado, se definen a continuación, qué variables intervienen en cada uno de estos elementos en la Tabla 1.

Tabla 1. Resumen de las variables principales

Pila del producto
Identificador
Prioridad
Descripción
Estimación del esfuerzo
Observaciones
Criterios de validación
Persona Asignada
Número de sprint en el que se realiza
Módulo del sistema al que pertenece
Tamaño de la pila del producto
Tiempo empleado en la elaboración de la pila del producto
Pila del sprint
Elemento de la pila del producto al que pertenece
Persona responsable de cada tarea
Estado en el que se encuentra cada tarea
Esfuerzo estimado
Sprint
Tamaño de la pila del sprint
Rendimiento del equipo en el último sprint (Velocidad)
Duración del sprint
Cantidad de sprints
Reunión de planificación
Duración de la reunión de planificación
Fecha de la reunión
Objetivo del sprint
Fecha de la reunión de revisión del sprint
Scrum diario
Duración del scrum diario
Reunión de revisión del sprint
Fecha para la reunión de planificación del siguiente sprint
Equipo de desarrollo
Tamaño del equipo
Estimación ágil
Velocidad

Luego de haber identificado las variables presentes en los proyectos de desarrollo de software Scrum, el siguiente paso fue analizar distintos trabajos de investigación referidos al proceso Scrum, en el campo de la simulación. De estas investigaciones, se determinaron cuáles son las variables más utilizadas por este tipo de

estudios, tomando como referencia las variables ya identificadas en la sección anterior. Las variables que se presentaron con mayor frecuencia son las que se consideran para el desarrollo de este trabajo. Los trabajos de investigación utilizados fueron [8], [18], [19], [20], [21], las variables presentes se ven en la Tabla 2:

Tabla 2. Variables en estudios de simulación.

Variable	Frec. de aparición
Cantidad de sprints	3
Duración de la reunión de planificación	2
Duración de la reunión de retrospectiva	2
Duración de los scrum diarios	2
Duración del sprint	5
Esfuerzo estimado de las tareas	1
Prioridad de las tareas	1
Puntos de las historias de usuario por sprint	1
Tiempo empleado en la elaboración de la pila del producto	2
Tamaño de la pila del producto	4
Tamaño de la pila del sprint	3
Tamaño del equipo	4
Velocidad	3

De las cuales, se calculó el porcentaje de aparición, considerándose como más importantes, a aquellas variables que tuvieran una presencia mayor o igual al 50% en las investigaciones analizadas. La siguiente tabla muestra cuáles son las variables con una mayor tasa de ocurrencia:

Tabla 3. Variables con mayor ocurrencia.

Variable	Porcentaje de aparición
Cantidad de sprints	50%
Duración del sprint	83,33%
Tamaño de la pila del producto	66,67%
Tamaño de la pila del sprint	50%
Tamaño del equipo	66,67%
Velocidad	50%

En base a la información reflejada por la tabla 3, se deciden utilizar las variables: **Cantidad de Sprints**, **Duración del Sprint**, **Tamaño de la Pila del Producto**, **Tamaño de la Pila del Sprint**, **Tamaño del Equipo** y **Velocidad** como variables fundamentales que deberán estar presentes en este trabajo.

4. Formación de Recursos Humanos

El equipo de trabajo se encuentra formado por cuatro investigadores con distintos niveles de posgrado, un Doctor en Ciencias Informáticas y Magister en Redes de Datos; un Doctor en Tecnologías de la Información y Comunicaciones Magister y Especialista en Ingeniería de Software; un Maestrando de Ingeniería de la Web; dos Maestrando en Redes de Datos y ocho estudiantes en período de realización de trabajos finales de grado en el contexto de las carreras de Licenciatura en Sistemas de Información y de Ingeniería en Informática de la UGD. Actualmente, el número de tesinas de grado aprobadas en el contexto de este proyecto, es de cinco, y otras tres en proceso de desarrollo. El número de tesis de maestría terminadas relacionadas con este proyecto es de una.

5. Bibliografía

- [1] Kent Beck, *Una Explicación de la Programación Extrema. Aceptar el Cambio*. España: Addison Wesley, 2002.
- [2] Borja López Yolanda, "Metodología Ágil de Desarrollo de Software – XP," , pp. 1-10.
- [3] "https://www.scrumalliance.org," Julio 2015.
- [4] Jerry Banks, "Introduction to Simulation ," in *Simulation Conference, 2000. Proceedings. Winter, 2000*, p. 7.
- [5] Breno Bernard Nicolau de França and Guilherme Horta Travassos, "Are We Prepared for Simulation Based Studies in Software Engineering Yet?," in *CLEI ELECTRONIC JOURNAL*, Montevideo: Centro Latinoamericano de Estudios en Informática, Abril 2013, p. 25.
- [6] Diego Alberto Godoy and Tamara Gisel Kasiak, "Modelo Dinámico de Simulación para la Gestión de Proyectos de Software Desarrollados con XP," in *XVIII Congreso Argentino de Ciencias de la Computación*, Octubre 2012, p. 10.
- [7] Godoy Diego Alberto, Belloni Edgardo A., Sosa Eduardo Omar, Kotynski Henry, and Benítez Juan de Dios, "Evaluación de alternativas de gestión en proyectos de software desarrollados con scrum utilizando dinámica de sistemas," in *XX Congreso Argentino de Ciencias de la Computación*, Buenos Aires, Octubre 2014, p. 10.
- [8] Rashmi Popli and Naresh Chauhan, "A Sprint-Point Based Estimation Technique," in *Information Systems and Computer Networks (ISCON), 2013 International Conference on*, Marzo 2013, p. 6.
- [9] Hugo A. Mitre Hernández, Edgar Ortega Martínez, and Cuauhtémoc Lemus Olalde, "Estimación y control de costos en métodos ágiles para desarrollo de software: un caso de estudio," in *Universidad Nacional Autónoma de México* , 2014, pp. 403-418.
- [10] Bhavani Seetharaman and Zulkefli Mansor, "The Development of Agile Cost Management Tool," in *International Conference on Electrical Engineering and Informatics (ICEEI)*, Diciembre 2015, p. 5.
- [11] Berenice Lencina, Yanina Medina, and Gladys N. Dapozo, "Aplicación para Estimar Costos en Proyectos de Software," in *Simpósio Argentino de Ingeniería de Software (ASSE 2016) - JAIIO 45 (Tres de Febrero, 2016)*, Septiembre 2016, pp. 181-192.
- [12] Cordero Carrasco and Raúl Jacinto, "Una Herramienta de Apoyo a la Estimación del Esfuerzo de Desarrollo de Software en Proyectos Pequeños," in *Repositorio Académico de la Universidad de Chile* , 2013, p. 114.
- [13] <https://confluence.atlassian.com/adminjiracloud/exporting-issues-776636787.html>, "Documentación Oficial de Atlassian Jira," Abril 2017.
- [14] <https://www.icescrum.com/documentation/advanced-exports/>, "Documentación Oficial de IceScrum," Abril 2017.
- [15] Trello Inc. (2017, Abril) Introducción a Trello. [Online]. <https://trello.com/guide/getting-started.html>
- [16] <http://help.trello.com/article/747-exporting-data-from-trello-1>, "Documentación Oficial de Trello," Junio 2016.
- [17] Luisanna Cocco, Katuscia Mannaro, and Giulio Concas, "A Model for Global Software Development with Cloud Platforms," in *38th Euromicro Conference on Software Engineering & Advanced Applications*, Enero 2012, pp. 446-452.
- [18] Godoy Diego Alberto, Belloni Edgardo A., Sosa Eduardo Omar, Kotynski Henry, and Benítez Juan de Dios, "Evaluación de Alternativas de Gestión en Proyectos de Software Desarrollados con Scrum Utilizando

- Dinámica de Sistemas," in *XX Congreso Argentino de Ciencias de la Computación*, Buenos Aires, Octubre 2014, p. 10.
- [19] Jaweria Sultana, "A STUDY ON APPLICABILITY OF THE SCRUM FRAMEWORK FOR LARGE SOFTWARE PROJECTS," Bachelor of Engineering, Tesis de Maestria 2015.
- [20] Isaac Griffith, Clemente Izurieta, Hanane Taffahi, and Claudio David , "A Simulation Study of Practical Methods for Technical Debt Management in Agile Software Development," in *Simulation Conference (WSC), 2014 Winter*, Savannah, GA, USA, 2014.
- [21] Ryushi Shiohama, Hironori Washizaki, Shin Kuboaki, Kazunori Sakamoto, and Yoshiaki Fukazawa, "Estimate of the Appropriate Iteration Length in Agile Development by Conducting Simulation," in *Agile Conference (AGILE)*, Dallas, TX, USA, 2012.

Movilidad y Gestión del Tráfico: Automatización y Aplicación de un Modelo de Calidad para Flotas Dinámicas en una Ciudad Inteligente

Giselle Cavallera (+); Carlos Salgado (+); Alberto Fernández Gil (*); Alberto Sánchez (+); Mario Peralta (+)
(+) Departamento de Informática - Facultad de Ciencias Físico-Matemáticas y Naturales - Universidad Nacional de San Luis
Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina
e-mail: giselle.cavallera@gmail.com, {csalgado, alfanego, mperalta}@unsl.edu.ar

(*) Centro para las Tecnologías Inteligentes de la Información y sus Aplicaciones (CETINIA)
Universidad Rey Juan Carlos
Madrid – España
e-mail: alberto.fernandez@urjc.es

Resumen

La gestión del tráfico es una de las áreas que implica mayores complicaciones en una ciudad [1] ya que afecta directamente el bienestar de los habitantes de la misma. En los últimos años han surgido nuevas necesidades que han requerido otro tipo de soluciones en cuanto a movilidad, y como resultado han surgido sistemas que promueven el uso compartido de vehículos [2], mediante el uso de tecnologías avanzadas.

Haciendo un análisis particular de los usuarios en interacción con las herramientas tecnológicas, es notable que junto a la madurez que han adquirido en el uso de las mismas, también se han incrementado sus requerimientos de calidad con respecto al software en general, en contextos reales de uso [3]. Esto implica que las aplicaciones software desarrolladas para los ciudadanos receptores, deben ser consistentes con esta realidad.

En este trabajo se propone la definición y desarrollo de una herramienta software que, mediante la utilización de un Modelo de Calidad, permite medir la calidad de una aplicación web-mobile de movilidad;

y brindar algunos indicadores que posibiliten el reconocimiento de aspectos débiles de la misma, para detectar dónde deben aplicarse mejoras.

Palabras clave:

Ciudades Inteligentes, Gestión del tráfico, Evaluación y automatización de métricas.

Contexto

El presente trabajo se enmarca en el proyecto de investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Además, se encuentra reconocido por el programa de Incentivos.

La labor se realiza en forma colaborativa con el grupo de investigación de Inteligencia Artificial (GIA), de la Universidad Rey Juan Carlos de Madrid-España. Dicho proyecto es la continuación de diversos proyectos de investigación sobre la gestión de flotas dinámicas y calidad aplicada a sistemas software, a través de los cuales se ha logrado un importante vínculo con la mencionada universidad internacional.

Introducción

Dos elementos aparecen como centrales en la mayoría de las descripciones de las ciudades inteligentes: los aspectos de transporte / logística, predominantemente desde el punto de vista de la sostenibilidad; y las nuevas tecnologías para facilitar la organización de actividades de una ciudad inteligente [4]. Éstas deben ser capaces de detectar los problemas de movilidad y tomar medidas en tiempo real para solucionar las posibles complicaciones. Además, la información que se obtiene dinámicamente de las tecnologías móviles tiene que ser facilitada a los ciudadanos para que puedan tomar sus propias decisiones [5].

Los nuevos sistemas de uso compartido de vehículos [1] son parte del desafío que plantean las Ciudades Inteligentes. Y ya que el objetivo de los mismos es la mejora de la movilidad humana y la reducción de costos, su análisis y evaluación es de gran utilidad para lograr su optimización.

El conjunto de vehículos que se usa de forma compartida para la movilidad de los ciudadanos conforma una flota. Algunas de ellas son totalmente controlables, pero otras apenas permiten un control limitado, como en el caso de, por ejemplo, el servicio público de bicicletas [2], el cual ya está implementado en grandes ciudades alrededor del mundo como París, Barcelona, Londres, Shanghai, Nueva York, entre otras [6].

La flota está disponible para todos los usuarios que utilizan sus vehículos a través de aplicaciones mobile o web. Ellos pueden reservar, tomar y dejar un vehículo cuando lo desean y dónde lo necesiten, siempre y cuando haya disponibilidad. Entonces, el ente que regula la flota, al usar la información sobre el estado dinámico de la misma, puede proveer y proponer mejores

opciones para las necesidades de transporte.

Cuando se habla de una aplicación de uso compartido de vehículos en una ciudad inteligente, se refiere a un sistema software con una gran cantidad de usuarios que está constantemente en uso. A modo de ejemplo, una de las plataformas de bicicletas compartidas más grandes del mundo, OFO [7], opera en más de doscientas ciudades y tiene más de doscientos mil usuarios. Esto implica que, los fallos vinculados a la falta de calidad de la aplicación podrían ser muy graves, generando caos en el transporte de las urbes.

Obtener un producto de alta calidad para la movilidad de los ciudadanos es esencial para impedir consecuencias negativas, ya que los beneficios de las herramientas software se ven opacados cuando el producto tecnológico no cumple con las condiciones de calidad requeridas para su uso, y corre el riesgo de generar un conflicto mayor, en lugar de ser un elemento valioso que contribuya a la mejora o resolución del problema [8].

Debido a que los involucrados e interesados en un producto software son muchos, y cada uno tiene diferentes objetivos, el factor clave para asegurar valor agregado a los mismos es una especificación y evaluación exhaustiva del software, lo cual puede lograrse definiendo las características de calidad necesarias y deseadas, asociadas a los objetivos y metas de los stakeholders [9] y efectuando la medición. En palabras de Peter Drucker: "Lo que no se puede medir, no se puede controlar; lo que no se puede controlar no se puede gestionar; lo que no se puede gestionar, no se puede mejorar"[10].

Como consecuencia de lo expuesto, la herramienta que se ha empezado a desarrollar se basa en un Modelo de Calidad cuyo pilar es la norma de calidad de producto ISO 25000 [9], el cual se ha comenzado a definir específicamente para una aplicación mobile de gestión de

tráfico en el contexto de las ciudades inteligentes.

Los resultados de la medición y evaluación de la calidad aportan al mejoramiento del software que se desarrolla y, consecuentemente, a la optimización de las aplicaciones y sistemas y sus ámbitos de aplicación.

Líneas de Investigación, Desarrollo e Innovación

De acuerdo a lo detallado previamente, en esta línea de investigación se trabaja en la definición y construcción de una herramienta software que permite la recolección de datos desde el dominio del problema. Puntualmente la herramienta recolecta datos/información en la ciudad de Madrid, España.

Los datos/información son parámetros requeridos por las métricas e indicadores del Modelo de Calidad utilizado, y el cálculo de los mismos. Esto permite confeccionar un reporte que proporciona un panorama de la situación real y que sirve de guía para llevar a cabo los cambios pertinentes en el sistema evaluado, junto con algunas recomendaciones y sugerencias.

El caso de estudio sobre el que se está trabajando es una aplicación de bicicletas compartidas llamada Ecobike [11], desarrollada en un trabajo de fin de grado de alumnos en Ingeniería Informática de la Universidad Rey Juan Carlos, Madrid.

Cabe destacar que un medio de transporte de rápido crecimiento en Europa son las bicicletas. Es así que muchos países están implementando sistemas para compartir bicicletas [12]. Por tanto, la medición de la calidad de una aplicación de este tipo, es de vital importancia para los ciudadanos.

En primer lugar, se han elegido las métricas del Modelo que pueden automatizarse.

La aplicación tiene definidos servicios REST (servicios que cumplen con la

arquitectura *Representational State Transfer*) que pueden ser consultados por terceros. Mediante herramientas que permiten conectarse a estos servicios, se obtienen los datos necesarios para hacer la medición.

Para medir la Eficiencia de Desempeño de la aplicación, por ejemplo, se seleccionó la subcaracterística *Comportamiento Temporal*. Una de las métricas elegida para ello ha sido el **Tiempo de Respuesta**, que establece cuánto tiempo transcurre antes que el sistema responda a una operación específica. Para efectuar el cálculo, se elige una función de la aplicación: reservar una bicicleta. A través de la herramienta que permite la comunicación con la API REST, se llama a la función y se obtiene el tiempo que llevó. Para obtener un promedio, se ejecuta la misma n veces y se calcula el mismo.

Para medir la Fiabilidad, se seleccionó la subcaracterística *Madurez*. Una de las métricas elegida para dicha medición ha sido la **Densidad de fallas**, que establece cuántas fallas fueron detectadas durante un periodo de prueba definido. Se ejecutan algunas funciones al azar por un determinado tiempo, se cuentan las fallas detectadas y se calcula la densidad:

$$X = A / B$$

Donde:

A: Número de fallas detectadas.

B: Tamaño del producto.

La herramienta usada permite obtener un archivo con todos los datos de la ejecución. Desde ahí se toman los fallos, y luego se calcula la densidad.

Los datos obtenidos se procesan mediante la aplicación de las métricas seleccionadas, y luego se generan los indicadores también definidos para obtener valores relevantes de análisis del software, los cuales permiten tomar cursos de acción en la optimización de la aplicación, y como consecuencia, impactar en el mejoramiento de la gestión de la flota de bicicletas compartidas.

Resultados y Objetivos

Desde el punto de vista de la Ingeniería de Software, una de las principales características que tiene que tener una aplicación software para ser exitosa entre los usuarios es que sea de calidad [6]. Para ello, se requiere generar valores, descriptores, indicadores o algún otro mecanismo de comparación. Por tanto, para garantizar la calidad es necesario llevar a cabo un proceso de medición del software [13], sin ignorar que el software móvil tiene sus características específicas para dispositivos móviles, por lo que los modelos y métodos correspondientes deben ajustarse a su ámbito móvil [14].

Los dispositivos junto con las aplicaciones móviles son utilizados en un cierto contexto, donde las características del mismo cambian continuamente. Los usuarios de éstos tienen ciertas particularidades, diferentes objetivos, realizan diferentes tareas; también se manipula en diferentes entornos físicos y sociales. Todos estos factores y otros, influyen en la forma de uso de una aplicación. Conviene aclarar entonces, que no se analiza la calidad en el entorno real de uso, en tanto que la aplicación y servicios se despliegan sobre una plataforma de simulación, y sería imposible tener en cuenta todas las variantes.

Como objetivo general, y como ya se ha mencionado en párrafos anteriores, se persigue concretar la medición de la calidad de la aplicación en estudio, cuya proyección de uso es muy amplia y cumple con la importante misión de aplacar los problemas recurrentes en la gestión del tráfico. Entre los objetivos específicos, que están en curso, se encuentra la construcción del módulo software para la recolección de datos requeridos por las métricas e indicadores, y el cálculo de los mismos. También, la utilización de éstos de forma adecuada para la creación de un informe relevante que proporcione un panorama que sirva

de guía para llevar a cabo los cambios pertinentes.

El mayor objetivo, de proyección a largo plazo, es el de contribuir a la optimización de la gestión del tráfico de una ciudad inteligente, de forma tal que el uso de productos software no represente una amenaza en la implementación de soluciones, sino que éstos puedan garantizar que son el elemento principal a través del cual las urbes logran continuar hacia un desarrollo sustentable.

Formación de Recursos Humanos

Bajo esta línea, en el grupo de investigación, se están desarrollando dos tesis de maestría en Calidad de Software. En una de ella se está definiendo/actualizando el modelo de calidad que cumple con los requerimientos de una aplicación software de movilidad para una Ciudad Inteligente. La otra consiste en la definición e implementación de una herramienta para la medición y evaluación de software de gestión de tráfico. Además, los investigadores Alberto Fernández y Holger Billhardt se encuentran trabajando sobre el proyecto de Coordinación Dinámica de Flotas Abiertas en Entornos Urbanos (SURF) [2], con inicio en el año 2016 y fin en el corriente año.

También se están llevando a cabo trabajos de grado con relación a la temática por alumnos de la carrera de Ingeniería Informática de la Universidad Rey Juan Carlos, Madrid.

Referencias

- [1] O. Mont. Institutionalisation of sustainable consumption patterns based on shared use. *Ecological economics*, 50(1):135– 153, 2004. The International Institute for Industrial Environmental Economics at Lund University.
- [2] Alberto Fernandez, Holger Billhardt. “Dynamic Co-ordination of Open Fleets in Urban Environments”

- (SURF), Centre for Intelligent Information Technologies (CETINIA). 2016-2018.
<http://sedici.unlp.edu.ar/handle/10915/4082>
- [4] Gary Graham y Lihong Zhang. Smart cities and digital technologies: the case of bike-sharing systems, 2015. <http://eprints.whiterose.ac.uk/84589/1/smart%20cities%20and%20bike-sharing%202015-1.pdf>
- [5] Cruzado, Juan Marín, “Evolución inteligente de la movilidad urbana”, 2017. <http://exeforum.biz/jornadas/smartmobility/kapsch.pdf>
- [6] Kabra, Ashish and Belavina, Elena and Girotra, Karan, Bike-Share Systems: Accessibility and Availability, 2015. Chicago Booth Research Paper No. 15-04.
- [7] www.ofo.com/us/en
- [8] Sommerville, Ian. “Ingeniería de Software Séptima edición”, Ed: Pearson Addison Wesley, 2005.
- [3] Covella, Guillermo Juan, “Medición y Evaluación de Calidad en Uso de Aplicaciones Web”, 2005.
- [9] ISO/IEC 25010:2011. Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models.
- [10] Drucker, P., “Management, W. H.”. Heinemann. 1975.
- [11] Ecobike Solutions: Gómez Pérez, Manuel y López González, Julio. Gestión de sistemas de alquiler de bicicletas. Trabajo Fin de Grado. URJC. 2016.
- [12] www.embers.city/new-ways-move-car-bike-sharing/
- [13] Fenton, N.E. y Pfleeger, S.L., Software metrics. A rigorous and practical approach, PWS Pub., 1997.
- [14] Dominik Franke, Stefan Kowalewski y Carsten Weise. A Mobile Software Quality Model, 2012. 12th International Conference on Quality Software. DOI: 10.1109/QSIC.2012.49

Hacia un Modelo de Calidad de Software como soporte a los procesos de licitación en el ámbito estatal.

Saldarini Javier*, Carrizo Claudio*, Armando Silvana*, Trasmontana Julio*, Salgado Carlos†, Sanchez Alberto†, Peralta Mario†

*Grupo de I+D Calidad de Software - Facultad Regional San Francisco Universidad Tecnológica Nacional
Av. de la Universidad 501 - San Francisco - Córdoba - Tel. 03564-421147
{saldarinijavier, jcarrizo77, silvana.armando, julio.trasmontana}@gmail.com

†Departamento de Informática Facultad de Ciencias Físico-Matemáticas y Naturales Universidad Nacional de San Luis Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina
e-mail: {csalgado, asanchez, mperalta}@unsl.edu.ar

RESUMEN

La modernización del Estado en todos sus estamentos es un tema que se ha instalado en estos últimos tiempos, en pos de la construcción de una gestión más transparente, eficiente, eficaz y al servicio de los ciudadanos. Esta modernización, no sólo se plantea a nivel del Estado Nacional, sino, que también alcanza a los Estados Provinciales y Municipales, que necesitan poner a disposición herramientas que permitan dirigirlos hacia una gestión orientada a servicios o lo que se conoce como Municipios 2.0.

Uno de los aspectos que se plantean en el ámbito de la Administración Pública Municipal, está relacionado con que habitualmente éstas no cuentan con recursos técnicos y/o metodológicos que les permitan realizar una toma de decisiones objetiva al momento de afrontar el proceso de adquisición de un producto software, dado que éste, no es un bien que estos tipos de organizaciones estén habituadas a adquirir.

Esta línea de I+D tiene como objetivo fomentar el desarrollo de modelos, métodos y/o nuevas herramientas, o adaptación de las existentes a través de la construcción de Modelos de Calidad para productos de software, a través de una metodología apropiada y tomando como base el catálogo de características de calidad propuestas en la norma de calidad ISO/IEC 25010.

La aplicación de estos modelos servirán como instrumentos para el soporte a la toma de decisiones al momento de la adquisición de productos de software, en el ámbito de las Administraciones Públicas Municipales.

Palabras Clave: *Calidad de Software, IQMC, ISO 25010, Modelos de Calidad.*

CONTEXTO

La presente línea de I+D se enmarca en el Proyecto de Investigación: *Gestión de Proyectos de Software: Los Modelos de Calidad como Soporte a los Procesos y Productos Software.*

Este proyecto fue elaborado y es coordinado en conjunto entre los integrantes del Laboratorio de Calidad e Ingeniería de Software (LaCIS) – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis; y, por otra parte, el Grupo de I+D Calidad de Software perteneciente a la Facultad Regional San Francisco, Universidad Tecnológica Nacional.

La evaluación, homologación y financiamiento del proyecto estuvo a cargo de la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional, el mismo es reconocido bajo el código: IAN4895.

1. INTRODUCCIÓN

En toda organización se desarrollan e implantan instrumentos y procedimientos

que posibilitan las funciones operativas y de toma de decisiones. Estos instrumentos pueden ser o no automatizados y cumplen funciones atendiendo a uno o varios procesos organizacionales y suelen estar definidos de manera formal o informal.

La Administración Pública Nacional (APN) es una organización basada en una administración compleja, con normas y reglas establecidas, esto aplica para todo organismo centralizado y/o descentralizado del Estado Nacional, y a partir de ello, los Gobiernos conforman una serie de estructuras organizativas/administrativas que permiten la organización del trabajo, la administración y el control según se muestra en [1].

En este marco se encuentran los Gobiernos Locales, mejor conocidos como las Municipalidades y denominadas técnicamente como las Administraciones Públicas Municipales (APM), articuladas con la APN y con la Administración Pública Provincial correspondiente.

Es innegable que en la última década los avances tecnológicos en el área de las Tecnologías de la Información y las Comunicaciones (TIC's) han cambiado la forma de hacer las cosas en las organizaciones, en este sentido la APN no es ajena a ello. Según se puede observar en [2] (GOBIERNO DIGITAL - Hitos significativos y evolución normativa período 1997-2015) el Estado Nacional Argentino ha venido desarrollando normativas que impulsan la modernización del estado nacional en pos de una gestión más eficaz y eficiente, también podemos ver el decreto 434/2016 donde se impulsa el plan denominado "Plan de Modernización del Estado" [3].

Las APM están inmersas en el cambio tecnológico y en particular de los que tienen relación directa con las TIC's, entonces es, en este sentido, que el mencionado Plan de Modernización del

Estado [3] tiene entre sus objetivos constituir una Administración Pública al servicio del ciudadano en un marco de eficiencia, eficacia y calidad en la prestación de servicios, a partir del diseño de organizaciones flexibles orientadas a la gestión por resultados.

Visto este marco, se puede visualizar que las APM se encuentran articuladas dentro de este Plan con la APN, a la hora de avanzar o comenzar a trabajar en lo que se denomina Municipios 2.0 [4], éstos deberán afrontar procesos de desarrollo, adquisición e implementación de productos y/o plataformas de software que les permitan dar el salto cuali y cuantitativo en lo que respecta a tecnologías TIC's aplicadas a la gestión.

En este marco de modernización del estado, cuando una APM tenga que afrontar el proceso de adquisición de un producto software, tal vez no cuente con los recursos técnicos y/o metodológicos, como para la adquisición de otros tipos bienes, que le permita realizar el pliego licitatorio con detalles técnicos y precisos respecto de:

- Las especificaciones funcionales y no funcionales del producto software.
- Determinación de las partes interesadas.
- Especificación de la calidad necesaria.
- Realización de la evaluación de las alternativas de los productos software ofrecidos.
- Ejecución de un control sobre el proveedor en el proceso de implementación del software.

El objetivo principal del presente trabajo consiste en construir un modelo de calidad mixto para evaluar productos software, lo que permitirá dotar al pliego licitatorio de un instrumento formal y objetivo a la hora de licitar, evaluar, implementar y controlar la implementación de un

producto software en el ámbito de la APM, tomando como punto de partida la perspectiva del adquirente del mismo.

Los Modelos de Calidad (MC), son instrumentos o artefactos específicamente diseñados y construidos para soportar evaluación y selección de componentes de software. Permiten la definición estructurada de criterios de evaluación, la especificación de requerimientos, la descripción de componentes en relación a ellos y la identificación de desajustes de manera sistemática facilitando el proceso de evaluación y selección del software [5]. Según Carvallo en [6], las propuestas existentes de modelos de calidad se pueden clasificar según si tienen un enfoque de modelos de calidad fijos, a medida o mixtos.

Entre los modelos de calidad fijos se pueden observar los de McCall et al. (1997) [7], Boehm et al. (1978) [8], entre otros.

Para los modelos de calidad a medida existen diversas propuestas de métodos para crearlos, como por ejemplo la del estándar IEEE 1061 [9].

Para el caso de los modelos de calidad mixtos se pueden destacar el ADEQUATE Horgan [10], el modelo de Gilb [11] y el modelo propuesto en el estándar ISO/IEC 9126-1 [12], este último es actualizado y reemplazado por el estándar ISO/IEC 25010:2011 [13].

La norma ISO/IEC 25010:11 [13] presenta un modelo de calidad del producto compuesto de ocho características (que luego se subdividen en treinta y dos subcaracterísticas) que se refieren a las propiedades estáticas del software y a las propiedades dinámicas del sistema informático. El modelo es aplicable tanto a sistemas informáticos como a productos software.

Los modelos pueden, por ejemplo, ser utilizados por desarrolladores, adquirentes, personal de aseguramiento,

de control de la calidad y evaluadores independientes, particularmente aquellos responsables de especificar y evaluar la calidad del producto software.

Cuando se habla de modelos de calidad se hace referencia a un conjunto de características y subcaracterísticas deseables por parte de algún actor (en este caso en particular, el adquirente, un Municipio) que son esperables que tenga un producto. Identificar y definir estas características permitirán tener los atributos necesarios para poder medir cada producto. En base a estos atributos se definen y/o reutilizan métricas ya sea de la bibliografía o definidas para realizar las mediciones.

Se detallan a continuación las 8 características que comprende el catálogo del modelo de calidad del producto según la norma ISO/IEC 25010:11 [13].

1-Adaptación funcional

2-Eficiencia del desempeño

3-Compatibilidad

4-Usabilidad

5-Confiabilidad

6-Seguridad

7-Capacidad de mantenimiento

8-Portabilidad

Coincidiendo con lo expresado en [6] (Carvallo J.P., et al. 2010 Capítulo 10), la construcción de modelos de calidad viene dificultada por distintas circunstancias relacionadas con: (1) el equipo que realiza la construcción del modelo, en el caso de que este equipo no tenga experiencia en la construcción de modelos de calidad o bien en el contexto del dominio del componente objeto; (2) el dominio para el que se construye el modelo, para el que en muchas ocasiones no existe una terminología común; (3) factores metodológicos, ya que es difícil conocer el nivel de profundidad hasta el que es necesario descomponer los modelos, y por tanto cuándo se puede decir que un modelo de calidad se ha finalizado.

También se plantea que la existencia de un método que proporcione unas pautas para la construcción de los modelos de calidad puede ayudar a paliar estas dificultades y presentan el método Individual Quality Model Construction (IQMC) que proporciona un conjunto de guías y técnicas para la identificación de los factores de calidad apropiados que deben ser incluidos en un modelo de calidad que permita analizar la calidad de componentes pertenecientes a un cierto dominio de software.

Este método (IQMC) consiste de 7 pasos que guían la construcción del modelo de calidad:

Paso 0. Estudio del ámbito del software.

Paso 1. Determinación de subcaracterísticas de calidad.

Paso 2. Refinamiento de la jerarquía de subcaracterísticas.

Paso 3. Refinamiento de subcaracterísticas en atributos.

Paso 4. Refinamiento de atributos derivados en básicos.

Paso 5. Establecimiento de relaciones entre factores de calidad.

Paso 6. Determinación de métricas para los atributos.

Si bien se presenta como una lista secuencial la construcción del modelo se puede llevar a cabo de manera iterada.

Tomando como base los modelos mixtos de calidad, la metodología IQMC para la construcción y el catálogo de referencia de características y subcaracterísticas de la norma ISO/IEC 25010, el objetivo es desarrollar modelos de calidad que sirvan como instrumentos objetivos y de soporte a las decisiones en los procesos licitatorios para la adquisición de productos software en el ámbito de las APM.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los principales ejes de esta línea de I+D están asociados a:

- Estudio de modelos conceptuales aplicados a la calidad de productos software.
- Estudio de normas de calidad aplicadas a productos software.
- Estudio de estándares y metodologías aplicadas a la construcción de Modelos de Calidad de productos.
- Construcción de modelos de calidad de productos software aplicables a diferentes ámbitos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los principales ejes de esta línea de I+D están asociados a:

- Estudio de modelos conceptuales aplicados a la calidad de productos software.
- Estudio de normas de calidad aplicadas a productos software
- Estudio de estándares y metodologías aplicadas a la construcción de Modelos de Calidad de productos.
- Construcción de modelos de calidad de productos software aplicables a diferentes ámbitos.

3. RESULTADOS OBTENIDOS/ESPERADOS

La línea de I+D presentada en este trabajo se está desarrollando tomando en cuenta los objetivos del Proyecto de I+D que la contiene.

De manera específica para esta línea se obtuvieron hasta la fecha los siguientes resultados:

- Caracterización del dominio y la problemática detectada, se indaga sobre los procesos en las APM respecto de las licitaciones en general y en particular las referidas a compras de software.

- Estudio de los modelos y normas de calidad aplicados a productos de software.
- Estudio de metodologías o métodos que guíen la construcción de modelos de calidad.
- Seleccionar un Municipio en el cual llevar a cabo un estudio de caso o proyecto experimental para poder poner a prueba el modelo conceptual construido.

Los resultados esperados son:

- Concluir con la definición del modelo conceptual de calidad aplicado a productos software a través de la metodología y norma de calidad seleccionada.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de I+D el equipo de trabajo está conformado por 3 Docentes Investigadores, 2 Becarios de Grado y 4 Tesistas de posgrado.

- Tesistas de posgrados:
 - Ing. Silvana Armando
 - Ing. Julio Trasmontana
 - Ing. Claudio Carrizo
 - Ing. Javier Saldarini
- Becarios de Grado: A designar en la convocatoria a becas 2018 de la Universidad Tecnológica Nacional

5. BIBLIOGRAFÍA

[1]. Mapa del Estado Nacional de la República Argentina. Recuperado en 2018:<http://mapadelestado.modernizacion.gob.ar/>

[2]. GOBIERNO DIGITAL - Hitos significativos y evolución normativa (período 1997-2015) Recuperado en 2018:
<http://www.inap.modernizacion.gob.ar/>

[3]. Decreto 434/2016 Plan de Modernización del Estado de la República Argentina

[4]. Gestión pública municipal. Recuperado en 2018:
<https://www.argentina.gob.ar/inap/altadireccionpublica/desarrollolocal>

[5]. J. Bermeo Conto, M. Sánchez, J. J. Maldonado, and J.P. Carvallo “Modelos de Calidad de Software en la Práctica: Mejorando su Construcción con el Soporte de Modelos Conceptuales”, CEDIA 1-abr-2016.

[6]. C. Calero, M. Piattini, M. Moraga, Calidad del producto y proceso software: Ra-Ma, 2010

[7]. MCCALL, J.A., RICHARDS, P.K. y WALTERS, G.F. (1977) “Factors in Software Quality”. RADC TR-77-369, Vols I, II, III, US Rome Air Development Center Reports NTIS AD/A-049.

[8]. BOEHM, B.W., BROWN, J.R., KASPAR, H., LIPOW, M., MACLEOD, G.J. y MERRITT, M.J. (1978). “Characteristics of Software Quality”. North Holland Publishing Company.

[9]. INSTITUTE OF ELECTRICAL ELECTRONIC ENGINEERING (1998). IEEE Std 1061-1998 IEEE

[10]. HORGAN, G., KHADDAJ, S. y FORTE, P. (1999) “Anessential Views Modelfor Software Quality Assurance”. En: Project Control for Software Quality, Shaker Publishing.

[11]. GILB, T. (1988). “Principles of Software Engineering Management”. Addison Wesley.

[12]. ISO/IEC 9126-1. (2001). Software engineering — Product quality — Part 1: Quality model.

[13]. ISO/IEC 25010:2011.

Smart City: Un modelo de calidad mixto para software responsable de la gestión de la movilidad urbana

Alejandro Ramón Rivoira⁽⁺⁾; Alberto Sánchez⁽⁺⁾; Alberto Fernández^(*); Carlos Salgado⁽⁺⁾; Mario Peralta⁽⁺⁾

⁽⁺⁾ Departamento de Informática - Facultad de Ciencias Físico-Matemáticas y Naturales - Universidad Nacional de San Luis
Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina
e-mail: rivoira.ale@gmail.com, {alfanego, csalgado, mperalta}@unsl.edu.ar

^(*) Centro para las Tecnologías Inteligentes de la Información y sus Aplicaciones (CETINIA)
Universidad Rey Juan Carlos
Madrid – España
e-mail: alberto.fernandez@urjc.es

Resumen

Las Smart Cities son ciudades que, por medio de las aplicaciones de la tecnología en sus diferentes ámbitos, se transforman en localidades más eficientes en el uso de sus recursos, ahorrando energía, mejorando los servicios entregados y promoviendo un desarrollo sustentable.

Las tecnologías de la información y de la comunicación son el eje central de las mismas, por tal motivo deben ofrecer garantía de calidad.

La calidad de software puede ser entendida como el grado con el cual el usuario percibe que el software satisface sus expectativas [1]. De esta definición se deduce que hay expectativas o deseos que él no percibe y que pueden acarrear un problema inusitado.

En este artículo se presenta una línea de investigación basada en la creación de un modelo de calidad mixto, acompañado de métricas e indicadores, donde se determinan y evalúan las características más relevantes de dos agentes software importantes para la gestión de la movilidad y el buen funcionamiento del tráfico en una ciudad inteligente; proveyendo valores reales de cumplimiento de dichas características. Ambos software son responsables de la gestión de coordinación dinámica de flotas abiertas, específicamente bicicletas.

Palabras Clave: Smart City, Movilidad, Coordinación Dinámica de Flotas Abiertas, Modelo de Calidad, ISO/IEC 25010, Métricas de Calidad.

Contexto

Esta propuesta está contextualizada en el trabajo colaborativo entre dos grupos de investigación: Los integrantes del Laboratorio de Calidad e Ingeniería de Software (LaCIS), Universidad Nacional de San Luis; y, por otra parte, el grupo de investigación interviniente, Inteligencia Artificial (GIA), de la Universidad Rey Juan Carlos de Madrid-España.

El trabajo se realiza en la Universidad Nacional de San Luis, en el Laboratorio de Calidad e Ingeniería de Software (LaCIS), dentro del contexto del Proyecto de Investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

Introducción

El aumento de la densidad de la población, generado por el crecimiento poblacional, y la alta concentración de la población en centros urbanos es un hecho generalizado en el mundo. Esta dinámica, crea la necesidad de pensar, prioritariamente, en modelos de desarrollo que organicen los procesos dentro de las ciudades por medio de sistemas que promuevan el uso eficiente de los recursos y, además, potencialicen la actividad económica y promuevan el desarrollo social [2].

El concepto Smart City (ciudades inteligentes) nace como idea global de la gestión de los recursos de una ciudad dirigidos a mejorar la calidad de vida de los ciudadanos. También se puede describir como aquella ciudad que aplica las tecnologías de la información y de la comunicación (TIC) con el objetivo de proveerla de una infraestructura que garantice:

- Un desarrollo sostenible.
- Un incremento de la calidad de vida de los ciudadanos.
- Una mayor eficacia de los recursos disponibles.
- Una participación ciudadana activa.

Por lo tanto, son ciudades que son sostenibles económica, social y medioambientalmente [3].

Para determinar si una ciudad es o no inteligente, generalmente, se toma como punto de partida un modelo en torno a seis características: *Economía, Gestión de gobierno, Ciudadanía, Entorno, Calidad de vida y Movilidad*. A partir de este modelo, se han desarrollado objetivos agrupados en una serie de servicios para los que se han descrito las principales tecnologías que disponemos en la actualidad. De esta forma, se obtuvo una serie de plataformas de servicios, que se deben integrar y conectar entre sí, facilitando que el ciudadano forme parte activa del proceso de gestión de su ciudad [4]. En otras palabras, se busca modernizar la gestión de las ciudades, fomentando una mayor interacción entre las instituciones, los ciudadanos y los sistemas informáticos [5].

Una Ciudad Inteligente utiliza TIC y otros medios para mejorar no solo la toma de decisiones, sino también la eficiencia de las operaciones, los servicios urbanos y su competitividad, a la vez que se garantiza la atención a las necesidades de las generaciones actuales y futuras en relación con los aspectos económicos, sociales y medioambientales [6]. A través de las TIC se integran las distintas áreas utilizando redes de comunicación de banda ancha, computación en la nube, dispositivos inteligentes móviles, programas de análisis y sensores. Este conjunto de recursos digitales capta datos generados por diferentes agentes (personas o dispositivos), procesa esos datos generando informaciones y permite construir y aplicar ese conocimiento para ofrecer mayor calidad de vida y beneficios a sus ciudadanos [7].

Si centramos el análisis solo en la *movilidad*, como uno de los factores importantes dentro de una Ciudad Inteligente, es imprescindible destacar que ante el reciente incremento de la preocupación por el medio ambiente y la mejora de la calidad de vida en el mundo entero, así como el interés de reducir factores que afectan al cambio climático, surge el uso de la bicicleta como una alternativa de bajo costo y no contaminante a los medios de transporte convencionales. Por lo tanto, poder garantizar que el uso de aplicaciones, para la gestión de flotas abiertas, específicamente bicicletas, sea aceptable; transforma la infraestructura tradicional de la ciudad en un ecosistema vivo y sostenible, brindando beneficios a las personas que viven y trabajan en ella.

Las aplicaciones o productos de software, cuando son lanzados al mercado, se espera que tengan cierto grado de aceptación entre los usuarios finales. Ese grado va a depender de las características particulares que cada uno de ellos considere importante a la hora de evaluar. Desde el punto de vista de la Ingeniería de Software, la calidad es una de las principales características que tiene que tener una aplicación para ser exitosa. Resulta relevante para los desarrolladores de software

poder medir la calidad o realizar pruebas de la misma a las aplicaciones construidas, siendo un requisito indispensable para poder medir, saber qué es lo que hay que medir y cómo.

La calidad de software en el sentido más general se define como: “Proceso eficaz de software que se aplica de manera que crea un producto útil que proporciona valor medible a quienes lo producen y a quienes lo utilizan” [6, 7]. Es, también, una compleja combinación de factores, que varía entre diferentes aplicaciones. Autores como Pressman [8], McCall [9] y estándares como ISO 9126 [10], ISO 14598 [11], ISO 25000 [12], entre otros, han tratado de determinar y categorizar los factores que afectan a la calidad del software.

Se conoce como modelo de calidad a un conjunto de buenas prácticas para el ciclo de vida del software, enfocado en los procesos de gestión y desarrollo de proyectos [18].

A lo largo del tiempo se han desarrollado diferentes modelos para evaluar la calidad del software. Entre los modelos más conocidos pueden mencionarse el de McCall [9], Evans y Marciniak [13], FURPS [14], Piattini, García y Caballero [15], entre otros.

Relacionado con la calidad del producto, recientemente ha aparecido la familia de normas ISO/IEC 25000.

ISO/IEC 25000 constituye una serie de normas basadas en ISO/IEC 9126 y en ISO/IEC 14598 cuyo objetivo principal es guiar el desarrollo de los productos de software mediante la especificación de requisitos y evaluación de características de calidad [12].

El modelo de calidad representa la piedra angular en torno a la cual se establece el sistema para la evaluación de la calidad del producto. Cada modelo cuenta con características, subcaracterísticas y métricas asociadas. Las métricas son una medida cuantitativa del grado en que un sistema, componente o proceso posee un atributo dado [19]. Los indicadores, por su parte, son métricas o combinación de métricas que

proporcionan una visión profunda, del proceso, del proyecto o del producto software [20]. Las métricas e indicadores son activos claves de una organización que proveen datos e información útiles para los procesos de análisis [16]. A su vez, también, son un buen medio para entender, monitorizar, controlar, predecir y probar el desarrollo del software y los proyectos de mantenimiento [17].

El modelo de calidad del producto, definido por la ISO/IEC 25010 se encuentra compuesto por ocho características: Adecuación Funcional, Eficiencia de desempeño, Compatibilidad, Usabilidad, Fiabilidad, Seguridad, Mantenibilidad y Portabilidad.

Líneas de Investigación y Desarrollo e Innovación

La investigación parte de la identificación del problema principal en el contexto definido. Una vez descrito el problema a solucionar se procedió a realizar un análisis, estudio y discusión de modelos de evaluación de calidad, estándares y metodologías a fin de que dicho modelo contemple la totalidad de las características a validar en los productos software intervinientes. Teniendo en cuenta el tipo de software (aplicación móvil, página web, aplicación web, API, etc.) y el rango del mercado al que se apunta.

En base a todo lo que se expresó anteriormente, se hizo la definición de un modelo de calidad mixto, contextualizado y adaptado, para medir y evaluar la calidad de EcoBike [21a] y BikeManager [21b] que son dos productos software responsables de la gestión de la movilidad, y por ende del tráfico, en una ciudad inteligente (específicamente Madrid). El modelo de calidad mencionado se denomina mixto porque, si bien tiene como base principal el Modelo ISO/IEC 25010, también se utilizaron indicadores o métricas del Modelo de Pressman, y algunas propias del dominio del problema.

Actualmente se está trabajando en el refinamiento de las métricas e indicadores de calidad para saber específicamente “qué” y “cómo” medir. Sin perder de vista que el uso de este modelo debe conducir a un conjunto de cambios en diversas partes del software para producir un notable aumento de su calidad.

Resultados obtenidos y Objetivos

La importancia, analizada como impacto social y económico, subyace quizás desde el punto de vista de la carencia. Desde nuestros entornos cotidianos el uso de sistemas informáticos genera implícitamente un sentido de confianza por un lado y automatización por otro. En primer término, confianza porque es recurrente revelar datos personales frente a cualquier aplicación que lo demande, sin conocer aún la procedencia o la garantía de seguridad que nos provee, por lo que confiamos a ciegas en que los desarrolladores de tal producto tienen en cuenta normas y estándares de calidad mínimos, sin conocer fehacientemente si es así. Cada descuido que el usuario se permite es lo que lleva al éxito de las grandes intrusiones, estafas y fraudes virtuales. Pero también, en segundo término, la automatización de uso tiene que ver con la rapidez con la que nos acostumbramos al manejo de una aplicación sin importar su dificultad. Hemos desarrollado un grado de inteligencia que nos lleva a interactuar de forma automática, desestimando si hay falta de coherencia o de sentido común en el uso o en la interfaz de la misma.

El modelo que se propone, intenta determinar el impacto de las deficiencias en un agente software dirigido a la gestión de la movilidad en ciudades inteligentes, teniendo como base la ciudad de Madrid. En relación a esto, se propondrán algunos elementos de solución que ayuden a crear un pensamiento consciente en las personas/usuarios.

Los sistemas evaluados (EcoBike [21a] y BikeManager [21b]) constan, principalmente,

de una aplicación web junto con un API REST y una aplicación móvil (Android) conectadas a un servidor común. Estas herramientas, en su conjunto permitirán simular el comportamiento de un sistema de alquiler de bicicletas en tiempo real.

El principal objetivo, por lo tanto, es contribuir a la mejora en la interacción entre los ciudadanos y los agentes software en una ciudad inteligente.

Se parte de los desarrollos realizados por grupos de investigación de la Universidad Rey Juan Carlos de Madrid y se tiene como objetivos fundamentales:

- Desarrollar la definición y validación de un modelo de calidad mixto para evaluar software de gestión de flotas de bicicletas en ciudades inteligentes. Partiendo del proyecto *Coordinación Dinámica de Flotas Abiertas en Entornos Urbanos* (Intelligent System for integrated and sustainable management of URban Fleets - SURF) [21].
- Proveer una base de información, a partir de las definiciones, mediciones y evaluaciones logradas, para automatizar la medición en tiempo real a través de un sistema informático.
- Brindar satisfacción a usuarios actuales, y generar atracción a nuevos usuarios a través de:
 - Protección y seguridad.
 - Adaptación y personalización.
 - Mejoras en usabilidad, rendimiento, eficiencia.

Los objetivos que se esperan alcanzar a largo plazo son los siguientes:

- Incremento de la productividad y satisfacción al trabajo de los profesionales afines al desarrollo de software.
- Mejora en la planeación eficaz de los agentes software.
- El acercamiento a cero defectos.

Finalmente, servirá para concientizar sobre la importancia de la calidad en la creación,

desarrollo y uso del software. Estando convencidos de que las mejoras propuestas permitirán un crecimiento sustancial en la vida cotidiana de una ciudad inteligente.

Formación de Recursos Humanos

Bajo esta línea, el grupo de investigación presentó un proyecto. Una Convocatoria Internacional de “Incorporación de investigadores Iberoamericanos” dentro del marco del Campus de Excelencia Internacional “Energía Inteligente” Programa CEISEP II. Convocatoria lanzada por la Universidad Rey Juan Carlos y la Universidad de Alcalá. Además, conjuntamente, se están desarrollando dos tesis de Maestría en Calidad de Software.

Referencias

- [1] IEEE Standard Glossary of Software Engineering Terminology, IEEE Standard 729-1983, IEEE, 1983.
- [2] Laura Liliana Moreno Herrera, Alejandro Gutiérrez Sánchez, *Ciudades Inteligentes: Oportunidades para generar soluciones sostenibles*, Estudios Sectoriales CINTEL - Proyectos TIC Innovadores, Colombia. 2012.
- [3] Endesa Educa. http://www.endesaeduca.com/Endesa_educ_a/recursos-interactivos/smart-city/. En el S.A. 2014.
- [4] Sergio Colado, Abelardo Gutiérrez, Carlos J. Vives, Eduardo Valencia. *SMART CITY: Hacia la gestión inteligente*. Editorial: S.A. MARCOMBO.
- [5] Fundación País Digital. *País Digital: Smart Cities*. <http://www.paisdigital.org/PD/smart-cities/>. Todos los derechos reservados.
- [6] Unión Internacional de Telecomunicaciones. *Grupo Temático sobre ciudades sostenibles e inteligentes*. 2014.
- [7] BID. Bouskela, M. Casseb, M. Bassi, S. De Luca, C. Facchina, M. *La ruta hacia las Smart cities: Migrando de una gestión tradicional a la ciudad inteligente*. 2016.
- [8] McCall J. A., PK Richards, GF Walters. “*Factors in Software Quality: Preliminary Handbook on Software Quality for an Acquisition Manager*”. Ed: DTIC Document. 1977.
- [9] McCall J.A, Richards P.K., Walters G.F. “Factors in Software Quality”. VOL I, II, III. NTIS. 1977.
- [10] ISO/IEC 9126-1:2001. Software Engineering - Software Product Quality - Part 1: Quality Model, Int’l Org. For Standardization, Geneva, 2001.
- [11] ISO/IEC 14598-1. Information Technology-Software Product Evaluation - Part 1: General overview, 1999.
- [12] ISO/IEC 25000. *SQuaRE - System and Software Quality Requirements and Evaluation*. <http://iso25000.com>
- [13] Michael W. Evans, John J. Marciniak. *Software Quality Assurance and Management*. Editorial: Wiley. 1987.
- [14] Grady, Robert; Caswell, Deborah. *Software Metrics: Establishing a Company-wide Program*. Prentice Hall. Hewlett-Packard (HP). 1987.
- [15] Piattini, M., García F., Caballero, I. “*Calidad de los Sistemas Informáticos*”. Editorial Alfaomega. 2007.
- [16] Olsina L., Pesotskaya E., Covella G., Dieser A. Bridging the Gap between Security/Risk Assessment and Quality Evaluation Methods. Publicada en ASSE Argentine Symposium on Software Engineering. 2012-2013.
- [17] Ian Sommerville. “*Ingeniería de Software Séptima edición*”. Ed: PEARSON Addison Wesley. 2005.
- [18] EcuRed. *Modelo de calidad: breves consideraciones*. https://www.ecured.cu/Modelo_de_calidad. La Habana, Cuba. 2005.
- [19] IEEE - Software Engineering Standards, Standard 610.12-1990, 1993.
- [20] Ragland, B. *Measure, Metric or Indicator: What is the Difference?* Ed: Crosstalk. 1995.
- [21] Alberto Fernandez, Holger Billhardt. Dynamic Co-ordination of Open Fleets in Urban Environments. SURF: Intelligent System for integrated and sustainable

management of URban Fleets. Centre for Intelligent Information Technologies (CETINIA). Se analizan específicamente 2 (dos) proyectos:

- a) Ecobike Solutions: Gómez Pérez, Manuel y López González, Julio. *Gestión de sistemas de alquiler de bicicletas*. Trabajo Fin de Grado. URJC. 2016.
- b) *Bikesmanager*: López Cerezo, Alejandro. *Sistema cliente/servidor para la gestión de parques de bicicletas*. Proyecto Fin de Carrera. URJC. 2016.

[Ingeniería de software para sistemas embebidos, requisitos en PYMEs y testing continuo

Irrazábal, Emanuel¹; Mascheroni, Maximiliano Agustín¹; Alonso, José Manuel¹, Vier, Sergio², Pereyra Coimbra, Rodrigo²

1: Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste

2: Departamento de Informática, Facultad de Ciencias Exactas Químicas y Naturales, Universidad Nacional de Misiones.

{eirrazabal, mascheroni}@exa.unne.edu.ar, alonso.651@gmail.com,
joroimbra@gmail.com sergiovier@gmail.com

Resumen

Esta línea de investigación aborda temas de ingeniería del software en sistemas tradicionales y su traslado a sistemas embebidos. Se espera la aplicación de la tecnología de desarrollo software en sistemas embebidos para el ámbito regional orientado a las entidades de ciencia y tecnología. En este sentido se está trabajando, por un lado, la construcción de un modelo de procedimientos para la gestión de requerimientos en entidades agrícolas con una cultura organizacional jerárquica, específicamente las entidades yerbateras del nordeste argentino. Y por otro lado en el desarrollo de sistemas embebidos de riego.

Finalmente se está trabajando en el desarrollo de procedimientos para la formalización del testing continuo en la disciplina de entrega continua lo cual facilitará los ensayos en los desarrollos ágiles.

Palabras clave: Ingeniería de software, requisitos, entrega continua, sistemas embebidos, riego

Contexto

La línea de Investigación y Desarrollo presentada en este trabajo corresponde al proyecto PI-F17-2017 “Análisis e implementación de tecnologías emergentes en sistemas computacionales de aplicación regional.”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2018-2021.

Asimismo, parte de la línea de investigación es realizada en el marco de la tesis de los maestrandos Sergio Fabián Vier y Rodrigo Pereyra Coimbra pertenecientes a la Maestría de Tecnologías de la Información Rs. 764/14 CS UNNE.

Introducción

Ingeniería de software para sistemas embebidos

Un sistema embebido es aquel sistema basado por lo general en un microprocesador, sensores y actuadores diseñado para realizar funciones dedicadas [1]. Y han cobrado gran importancia desde el punto de vista de los sistemas de información con el uso de plataformas tipo Arduino para el desarrollo rápido de prototipos [2].

En este sentido, el foco del grupo de investigación estará puesto en el desarrollo

de soluciones para entidades regionales, en particular para la agricultura y los grupos de investigación de la universidad. La actividad agrícola es una de las principales fuentes de producción de alimentos en el ámbito mundial. Las nuevas tecnologías de la computación y la comunicación ofrecen soluciones que pueden adaptarse a las necesidades de la agricultura para una producción más eficiente en el uso de recursos y por lo tanto más sustentable económica, ambiental y socialmente. Entre ellas se encuentran los desarrollos relacionados con la automatización del riego. Desde hace varios años se han implementado sistemas de riego que permiten un uso más racional del agua con niveles de desarrollo óptimos, en los cultivos donde se implementen [3][4][5][6][7]. Para los diferentes cultivos existen requerimientos particulares con los que lograr un óptimo desarrollo y productividad. Cada fase del desarrollo del cultivo requiere un mínimo de acumulación de temperatura, así como necesidades hídricas y nutricionales que varían a lo largo de cada fase. Esto hace indispensable tener un control asistido de estos parámetros [8]. Asimismo, el Instituto Nacional de Tecnología Agropecuaria (INTA) desarrolla diferentes líneas de investigación para mejorar el desarrollo de los cultivos [9]. Y una de sus necesidades es asegurar el control de las variables ambientales en los laboratorios e invernáculos. Por lo tanto, la inclusión de los sistemas de riego automatizados permite el uso eficiente del agua, es capaz de mejorar el rendimiento de los cultivos y es un insumo para el desarrollo de actividades de investigación agropecuaria.

Ahora bien, los esquemas de riego estáticos tradicionales no se ajustan adecuadamente a las necesidades hídricas de los cultivos en producción o de los ensayos en el ámbito de investigación [10]. El cambio en las temperaturas y las precipitaciones durante el día requieren, por lo tanto, un manejo hídrico no lineal.

De la misma manera, el desarrollo de tecnologías sustitutas en los laboratorios de investigación son una práctica corriente en los países en vías de desarrollo, por su alto costo y las barreras de mantenimiento [11][12]; esto especialmente en el interior del país. Asimismo, permite la personalización del equipamiento y la inclusión tecnologías como el diseño 3D [13] o los sistemas embebidos para la monitorización y el control de los procesos del laboratorio [14].

Ingeniería de Software para Requisitos y Entrega Continua

Tal y como lo describe la segunda ley de Lehman [15], el software necesita ser cada vez más complejo para satisfacer las necesidades de los usuarios. Esto, sumado a la importancia de detectar errores en la etapa de captura de requerimientos por sus costos [16], hace a la validación de los requerimientos una etapa crítica en el desarrollo de las aplicaciones [17].

A su vez, la mayoría de las Empresas Yerbateras que se encuentran en la provincia de Misiones tienen una estructura jerárquica tradicional, donde se identifican necesidades para gestión y seguimiento de tareas relacionadas a la Ingeniería de Software [18]. Estas empresas en su mayoría cuentan con un directorio que toma las decisiones, el cual conoce el negocio, pero no cuenta con el tiempo o la heurística suficientes para manifestar sus necesidades.

En este sentido, los desarrollos de metodologías ágiles representan un avance en la manera de construir sistemas; usando métodos, como, por ejemplo: el prototipado, las historias de usuario y los casos de uso. Esto hace posible la entrega temprana de valor, la respuesta rápida a los cambios y la colaboración constante del equipo de trabajo con los clientes [19][20].

Entonces, haciendo foco en la gestión de los requerimientos, es posible adaptar técnicas ágiles las cuales tendrán un fuerte impacto positivo [21][22], especialmente

en PYMEs con problemas al momento de la gestión de sus requerimientos.

Las herramientas y prácticas genéricas deben ser seleccionadas y adaptadas para la empresa, tamaño y dominio específico. Y así, reducir los problemas causados por el no alineamiento tales como requisitos implementados incorrectamente, retrasos y esfuerzos desperdiciados [25].

El Sistema de Información Universitaria (SIU) [26] inicia en 1996 desarrollando sistemas informáticos para la gestión de diversas áreas en instituciones que componen el sistema universitario nacional. El SIU cuenta con 54 proyectos de software activos [27]. Uno de los proyectos recientes y actualmente en desarrollo es el SIU-Araí [28]. Se trata de una plataforma integradora de servicios, donde cada uno de los sistemas desarrollados por el SIU es considerado un módulo. De esta forma, la plataforma SIU-Araí es concebida como un componente central en las soluciones que proporciona el SIU hacia las instituciones del sistema universitario argentino.

A la par, la industria del software está moviéndose hacia la adopción de un paradigma en el cual la funcionalidad del software está en continua evolución [29]. Y el SIU busca llevar adelante el desarrollo de las soluciones software con un enfoque evolutivo.

Esto es cubierto por el enfoque de la ingeniería del software conocido como Entrega Continua de Software, en inglés Continuous Delivery (CD). Este es un enfoque en el cual los equipos mantienen la producción de software en ciclos cortos de tiempo, asegurando que el producto pueda ser lanzado de manera fiable en cualquier momento [30]. El objetivo es poder lanzar a producción un producto software libre de defectos “*con solo apretar un botón*” [31].

Uno de los principales desafíos de este enfoque es la calidad del producto software resultante. Ésta puede disminuir, dado que, al realizarse los despliegues del sistema con mayor frecuencia, aparecen más

defectos en el producto. Por tanto, es esencial desarrollar un enfoque de priorización de los diferentes aspectos en la calidad del producto software, teniendo en cuenta la forma de trabajo actual de las empresas de desarrollo software.

Líneas de investigación y desarrollo

En la línea de Ingeniería de Software para Sistemas Embebidos se propone:

- Desarrollar soluciones de sistemas embebidos con microcontroladores que solucionen problemas regionales y apoyen a los grupos de investigación locales.

En la línea de Ingeniería de Software para los Requisitos y la Entrega Continua:

- Desarrollar un modelo de procedimientos para la gestión de requerimientos en PYMEs yerbateras a partir de técnicas ágiles.
- Diseñar un modelo de pruebas de software mediante la combinación de las técnicas existentes para su uso en el SIU.
- Desarrollar una herramienta que permita la ejecución de pruebas de manera continua basadas en el modelo diseñado.

Resultados obtenidos

El grupo de investigación es de reciente formación, por lo cual los resultados son preliminares y, en parte, se enumeran antecedentes llevados adelante en el marco de otros grupos de trabajo. A continuación se indican:

En la línea de Ingeniería de Software para Sistemas Embebidos:

- Se realizaron dos presentaciones a congresos especificando el desarrollo de un Planificador Embebido para la Gestión de Riego Automatizado (PEGRA). En el primer artículo se describe el diseño y la construcción del prototipo [32]. En el segundo artículo se

detalla el diseño con componentes de uso específico y mejoras al diseño [33].

En la línea de Ingeniería de Software para la Entrega Continua:

- En [32] se realizó una revisión sistemática de la literatura. En ella, se buscaron propuestas, técnicas, enfoques, métodos, herramientas y otro tipo de soluciones para afrontar los problemas de pruebas en entornos de desarrollo continuo. También se validó que las pruebas continuas son un elemento faltante en la entrega continua, y se analizaron las diferentes definiciones de la misma y los diferentes niveles y etapas de pruebas. Por último, se detectaron nuevos problemas aún no resueltos relacionados a las pruebas continuas.
- En [33] se llevó a cabo una revisión sistemática con el objetivo de identificar, analizar y sintetizar las técnicas, herramientas y desafíos encontrados en la literatura sobre pruebas de compatibilidad web. Los resultados indicaron que la técnica de prueba más elegida es el análisis visual. Asimismo, el principal desafío detectado es la detección de elementos variables.

Formación de recursos humanos

En el Grupo de Investigación en Innovación en Software y Sistemas Computacionales (GISSC) están involucrados 4 docentes investigadores, 1 becario de investigación de pregrado, 1 tesista de doctorado y 3 tesistas de maestría. Cinco alumnos de la carrera están realizando sus proyectos finales vinculado a estos temas.

Referencias

- [1] Heath, Steve. *Embedded systems design*. Elsevier, 2002.
- [2] Jamieson, Peter. "Arduino for teaching embedded systems. are computer scientists and engineering educators missing the boat?." Proc. FECS 2010 (2010): 289-294.
- [3] Martínez, Germán Arturo López, Moreno, Flor García y Fierro, Juan Bedoya. Modelo a escala de un sistema de riego automatizado, alimentado con energía solar fotovoltaica: nueva perspectiva para el desarrollo agroindustrial colombiano. 2014, Revista Tecnura, Vol. 17, págs. 33-47.
- [4] Castro Silva, Juan Antonio. Sistema de riego autónomo basado en la internet de las cosas. Neiva, Colombia: s.n., 2016.
- [5] Escalas Rodríguez, Gabriel. Diseño y desarrollo de un prototipo de riego automático controlado con Raspberry Pi y Arduino. Barcelona, España: s.n., 2014.
- [6] Escamilla Martínez, Fernando. Automatización y telecontrol de sistema de riego. Gandía, España: s.n., 2016.
- [7] Hernandez Garcia, C. Itzel. Diseño de un sistema automatizado de riego para el cultivo de tomate y pimiento bajo invernadero. Xalapa Enríquez, México: Diciembre de 2014.
- [8] An Intelligent Smart Irrigation System Using WSN and GPRS Module. Manimaran , P. . 6, 2016, International Journal of Applied Engineering Research, Vol. 11, págs. 3987-3992.
- [9] Instituto Nacional de Tecnología Agropecuaria. INTA. [En línea] 2013. [Citado el: 12 de Junio de 2017.]
- [10] Pereira, Luis Santos, y otros, y otros. Riego localizado. El riego y sus tecnologías. Albacete, España : Europa-América, 2010, págs. 229 - 231.
- [11] Pearce, Joshua M. Open-source lab: how to build your own hardware and reduce research costs. Newnes, 2013.
- [12] Lazalde, Alan, Jenny Torres, and David Vila-Viñas. "Hardware: ecosistemas de innovación y producción basados en hardware libre." Buen Conocer-FLOK Society. Modelos sostenibles y políticas públicas para una economía social del conocimiento común y abierto en el Ecuador. Asociación aLabs, 2015. 619-652.
- [13] Baden, Tom, et al. "Open Labware: 3-D printing your own lab equipment." PLoS biology 13.3 (2015): e1002086.
- [14] Pearce, Joshua M. "Building research equipment with free, open-source hardware." Science 337.6100 (2012): 1303-1304.

- [15] M. Lehman, "On Understanding Laws, Evolution and Conservation in the Large Program Life Cycle", *J. of Sys. and Software*, vol. 13, pp. 213-221, 1980.
- [16] S. Maalem and N. Zarour, "Challenge of validation in requirements engineering", *Journal of Innovation in Digital Ecosystems*, vol. 3, no.1, pp. 15-21, 2016.
- [17] G. Kotonya and I. Sommerville, *Requirements engineering*. Chichester: John Wiley & Sons, 1998.
- [18] Plan Estratégico para el Sector Yerbatero – Yerba Mate Argentina", *Yerbamateargentina.org.ar*, 2016.
- [19] A. Sillitti and G. Succi, *Requirements Engineering for Agile Methods*. In: Aurum A., Wohlin C. (eds) *Engineering and Managing Software Requirements*. Berlin, Heidelberg: Springer, 2005.
- [20] I. Inayat, S. Salim, S. Marczak, M. Daneva and S. Shamshirband, "A systematic literature review on agile requirements engineering practices and challenges", *Computers in Human Behavior*, 2014.
- [21] S. Dragicevic, S. Celar and L. Novak, "Use of Method for Elicitation, Documentation, and Validation of Software User Requirements (MEDoV) in Agile Software Development Projects", 2014 Sixth International Conference on Computational Intelligence, Communication Systems and Networks, 2014.
- [22] One, V. "State of agile development survey results", 2017.
- [23] E. Bjarnason, P. Runeson, M. Borg, M. Unterkalmsteiner, E. Engström, B. Regnell, G. Sabaliauskaite, A. Loconsole, T. Gorschek and R. Feldt, "Challenges and practices in aligning requirements with verification and validation: a case study of six companies", *Empirical Software Engineering*, vol. 19, no. 6, pp. 1809-1855, 2013.
- [24] «Sistema de Información Universitario», ¿Qué hacemos? [En línea]. Disponible en: <http://www.siu.edu.ar/que-hacemos/>. [Accedido: 24-jul-2017].
- [25] «Sistema de Información Universitario», *Nuestra gestión en números*. [En línea]. Disponible en: <http://www.siu.edu.ar/nuestra-gestion-en-numeros/>. [Accedido: 24-jul-2017].
- [26] «Documentación - SIU-Arai». [En línea]. Disponible en: <http://documentacion.siu.edu.ar/wiki/SIU-Arai>. [Accedido: 25-jul-2017].
- [27] J. Järvinen, T. Huomo, T. Mikkonen, y P. Tyrväinen, «From agile software development to mercury business», en *International Conference of Software Business*, 2014, pp. 58–71.
- [28] L. Chen, "Continuous Delivery: Huge Benefits, but Challenges Too" in *IEEE Software* 03/2015. V. 32(2).
- [29] J. Humble and D. Farley. "Continuous delivery: reliable software releases through build, test, and deployment automation", 1st ed. Boston, US: Pearson Education, 2010.
- [30] Alonso, JM; Ferrari, H; Sambrana, I; Irrazábal, E. Emanuel Irrazábal. Desarrollo de un Planificador embebido para la gestión de riego automático aplicado a invernáculos. Implementación en INTA El Sombrero. XXIII Congreso Argentino de Ciencias de la Computación CACIC 2017. ISBN: 978-987-4251-43-5 La Plata (Argentina). Fecha: Octubre de 2017
- [31] Alonso, JM; Ferrari, H; Sambrana, I; Irrazábal, E. Desarrollo de un Sistema de Gestión de Riego Automático Basado en la Humedad del Suelo. 5to Congreso Nacional de Ingeniería Informática / Sistemas de Información Aplicaciones Informáticas y de Sistemas de Información. Libro de Actas CONAISSI 2017. Santa Fe (Argentina). Fecha: 2 y 3 de Noviembre de 2017
- [32] Mascheroni, M.A., Irrazábal, E. (in press). Continuous Testing and solutions for testing problems in Continuous Delivery: A Systematic Literature Review. *Computación y Sistemas*.
- [33] Sabaren, L; Mascheroni, M.A.; Greiner, C; Irrazábal, E. Una Revisión Sistemática de la Literatura en Pruebas de Compatibilidad Web XXIII Congreso Argentino de Ciencias de la Computación CACIC 2017. ISBN: 978-987-4251-43-5. La Plata (Argentina). Fecha: Octubre de 2017

PROPONIENDO UN ENFOQUE INTEGRADOR PARA DISEÑAR Y EVALUAR INTERFACES DE USUARIO WEB

**Adriana MARTIN, Gabriela GAETAN, Viviana SALDAÑO, Claudia CARDOZO,
Silvia VILLAGRA, Alejandra CARRIZO, Fernando VARGAS**

Grupo de Investigación y Formación en Ingeniería de Software (GIFIS)
Instituto de Tecnología Aplicada (ITA)

Universidad Nacional de la Patagonia Austral, Unidad Académica Caleta Olivia (UNPA-
UACO)

{amartin// ggaetan// vivianas // svillagra// acarrizo}@uaco.unpa.edu.ar;
claudia_yoryi@yahoo.com.a; ferez_vargas_94@hotmail.com

RESUMEN

Desarrollar productos Web para satisfacer las necesidades y preferencias de grupos de usuario de interés, requiere no sólo considerar los aportes de todas las áreas comprometidas con el lado humano de la Web, tales como Usabilidad, Accesibilidad Web y Diseño Centrado en el Usuario (DCU), sino también, trabajar estos aportes de manera conjunta dentro de un enfoque integrador que permita reforzar y potenciar las fortalezas, minimizando las debilidades.

El Proyecto de Investigación (PI) que se presenta en este trabajo, estuvo dirigido a desarrollar productos Web que atiendan las expectativas y necesidades de grupos de usuarios de interés, proponiendo, aplicando y validando un enfoque integrador de técnicas y herramientas conceptuales y prácticas basadas en la eXperiencia de Usuario (UX), a los efectos de: (i) satisfacer a usuarios adultos mayores; (ii) incluir aportes de las áreas vinculadas al lado humano de la Web para contribuir al desarrollo de productos enfocados en mejorar la UX; y (iii) considerar la UX en el desarrollo Web móvil.

Palabras clave: *Enfoque Integrador | Experiencia de Usuario (UX) | Diseño y Evaluación | Usabilidad | Accesibilidad Web | Grupos de Usuarios Adultos Mayores.*

CONTEXTO

El Grupo de Investigación y Formación en Ingeniería de Software (GIFIS) perteneciente al Instituto de Tecnología Aplicada (ITA),

Universidad Nacional de la Patagonia Austral (UNPA), Unidad Académica Caleta Olivia (UACO), ha finalizado la ejecución del Proyecto de Investigación (PI) N° 29/B194, Período: 2016-2018, Denominado: “*Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web*”, dirigido por la Dra. Adriana Martín y codirigido por la Mg. Gabriela Gaetán. Para desarrollar la problemática propuesta por este PI, GIFIS cuenta con sólidos antecedentes y experiencia recabada desde 2010 a través de la ejecución de los siguientes Proyectos de Investigación: PI N° 29/B167, Período: 2014-2016, Denominado: “*Identificación, Desarrollo y Uso de Soluciones Web Centradas en el Usuario*” y, PI N° 29/B144, Período: 2012-2014, Denominado: “*Diseño y Evaluación de Portales Web*”.

En este contexto, y para alcanzar la concreción del PI N° 29/B194, hemos dirigido nuestros esfuerzos en proponer un enfoque que, integrando herramientas conceptuales y prácticas en un ciclo de diseño y evaluación de interfaces para productos Web, que permita considerar y mejorar la experiencia de grupos de usuarios adultos mayores. En el transcurso de la ejecución del PI, incorporamos al objetivo inicial la problemática de diseñar con UX para móviles, debido a que las personas experimentan día a día la necesidad de interactuar con todas sus aplicaciones a través de sus múltiples dispositivos. Es importante destacar, que los resultados alcanzados por el PI N° 29/B194, Período: 2016-2018, ya se están aplicando en el nuevo PI N° 29/B222, Período:

2018-2020, Denominado: “*Diseño y Evaluación de Experiencia de Usuario para Multi-Dispositivos*”, dirigido por la Dra. Adriana Martín y codirigido por la Mg. Gabriela Gaetán.

1. INTRODUCCIÓN

Es muy difícil imaginar un escenario donde la interacción dentro de las sociedades se lleve a cabo sin la intervención de la Web. La Web se ha convertido en el instrumento preferido y adoptado no sólo por la mayoría de las personas, sino que también por las organizaciones públicas y privadas, sin fines de lucro y de negocio. Esto es debido a que la Web constituye una herramienta muy poderosa a la hora de acercar sitios y aplicaciones destinados a proveer información, comunicación, productos y servicios a los ciudadanos. Sin embargo, aún no se ha tomado completa conciencia del rol que juegan las personas destinatarias de esta vertiginosa evolución tecnológica: las preferencias de los verdaderos usuarios de los sitios y aplicaciones desplegados en la Web.

Numerosas áreas del conocimiento aportan propuestas y herramientas para mejorar el tratamiento de las necesidades de los usuarios, tales como: Diseño Centrado en el Usuario (“User-Centered Design”), Interacción Hombre-Computadora (“Human-Computer Interaction”), Estrategias de Contenido Web (“Content Strategy”), Usabilidad y Accesibilidad Web, entre otras.

Particularmente, en los últimos años y relacionado con esta necesidad de identificar más completamente a los usuarios Web para poder reconocer sus expectativas respecto a un producto desarrollado para la Web, el término “experiencia de usuario” (UX: User eXperience), ha cobrado protagonismo en los ámbitos de investigación y de los negocios. El término UX evoca un espectro mucho más amplio para atender a la diversidad humana, que el propuesto por las áreas del conocimiento antes mencionadas desde sus nichos de trabajo y aportes focalizados.

Diseñar para la experiencia del usuario, implica interacción social y cultural, diseño con valor sensitivo e impacto emocional, es decir, cómo la experiencia de acceso e interacción con el producto Web incluye alegría, comunicación emocional, impresión senso-perceptivas en el

uso, etc., lo que se denomina en el diseño de interfaces de usuario “*look and feel*”.

Desde el Grupo de Investigación y Formación en Ingeniería de Software (GIFIS), se viene trabajando fuertemente en esta problemática desde 2012 y particularmente en los dos últimos años este trabajo ha permitido identificar valiosos aportes y recursos en el área, tales como [1][2][3][4][5][6][7][8], entre muchos otros.

La experiencia recabada durante varios años de investigación, permiten aseverar, que cuando se trata de mejorar la UX en la Web, no sólo la Usabilidad es un componente clave, sino que también lo son otros factores de calidad, tales como la Accesibilidad Web, que es un recurso indispensable para aspirar a alcanzar una justa consideración de la diversidad de las personas destinatarias (usuarios Web), sus necesidades y características distintivas y diferentes. Con esta motivación, PI N° 29/B194 se definieron los siguientes Objetivos:

Objetivo Direccional: “Desarrollar productos Web aplicando un enfoque integrador de técnicas y herramientas basadas en la experiencia del usuario (UX), que satisfaga expectativas y necesidades de perfiles de usuarios de interés.”

Objetivos Operacionales:

OO.1 Explorar requerimientos de información, comunicación y servicios de los usuarios pertenecientes al perfil de interés.

OO.2 Explorar necesidades de Usabilidad, Accesibilidad e interacción de los usuarios pertenecientes al grupo de interés.

OO.3 Seleccionar, extender y/o proponer técnicas y herramientas para satisfacer los requerimientos y necesidades de 1. y 2.

OO.4 Proponer enfoque integrador de las técnicas y herramientas de 3. que mejore la experiencia de los usuarios (UX) pertenecientes al grupo de interés.

OO.5 Aplicar el enfoque integrador propuesto en 4. en el desarrollo de productos Web que satisfagan expectativas y necesidades de los grupos de usuarios de interés.

OO.6 Validar los productos Web de 5. con los usuarios pertenecientes al grupo de interés.

Los integrantes de GIFIS han estado durante años participando activamente en el dictado de

cursos de capacitación a la población de los Adultos Mayores. La valiosa experiencia recabada en la identificación de las preferencias y necesidades de esta porción poblacional creciente, han sido un factor determinante para enfocar particularmente los esfuerzos en este grupo de usuarios de interés.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El Proyecto de Investigación (PI) N° 29/B194, Período: 2016-2018, denominado “*Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web*”, ejecutó un Plan de Actividades compuesto por 3 Líneas de Investigación, las cuales comparten como hilo conductor los Objetivos Operacionales:

LI.1: Usuarios Adultos Mayores

Desde 2013, GIFIS ha estado vinculado al dictado de los Cursos de Computación en el marco del Convenio UPAMI (UNPA-PAMI). Y a partir de 2016, está a cargo de la Elaboración de la Propuesta y Dictado de Cursos de Computación a nuestros Abuelos de la región. Este espacio de intercambio e interacción ofrece el marco adecuado para ejecutar la LI.1.

Actividades de la LI.1:

A1.1 Definir un marco de experimentación para el análisis de las necesidades de usuarios adultos mayores.

A1.2 Evaluar la aplicabilidad de técnicas y herramientas que permitan satisfacer necesidades de usuario adultos mayores.

A1.3 Desarrollar y validar productos Web dirigidos a satisfacer necesidades de usuarios adultos mayores.

LI.2 Experiencia de Usuario (UX)

Entre 2013-2015, GIFIS ha estado trabajando en la identificación de problemas en los sitios Web Universitarios para los usuarios ciegos y en la aplicación de patrones de accesibilidad para asistir a la solución de estos problemas. A partir de 2016, se han recibido pedidos de desarrollo de productos Web que satisfagan las necesidades de los usuarios pertenecientes a comunidades específicas. La necesidad de satisfacer las necesidades de los usuarios pertenecientes a dominios específicos de interés, ofrece el marco adecuado para ejecutar la LI.2.

Actividades de la LI.2:

A2.1 Definir un enfoque UX que considere propiedades de Accesibilidad.

A2.2 Desarrollar un producto Web que aplique el enfoque UX propuesto.

A2.3 Validar el producto Web considerando las necesidades de un grupo de usuarios de interés.

LI.3 Web Móvil

Desde 2016, y a los efectos de explorar las tendencias en el desarrollo de productos Web móviles (*tablets*, celulares, etc.) para diferentes grupos de usuarios de interés, considerando la UX, GIFIS ha decidido llevar a cabo la LI.3.

Actividades de la LI.3

A3.1 Identificar técnicas y herramientas para la Web móvil.

A3.2 Analizar las necesidades de grupos de usuarios de interés.

A3.3 Proponer un enfoque UX para desarrollo Web móvil que incluya la participación de los usuarios.

3. RESULTADOS OBTENIDOS/ ESPERADOS

En lo que se refiere a resultados obtenidos, GIFIS viene trabajando desde 2012 en el área del conocimiento relacionada con la problemática abordada por el presente PI, correspondiendo al período 2016-2018, la siguiente producción [9][10][11][12][13][14][15][16][17].

3.1. Resultados Esperados

Alineado a los objetivos (presentados en la Sección 1.), los objetivos fijados para el presente PI, fueron los siguientes:

1. Desarrollar prototipos/ productos Web que satisfagan necesidades de grupos de usuarios de interés.
2. Aplicar un enfoque UX integrador al desarrollo de los productos Web en 1.
3. Interactuar con los grupos de usuarios de interés que permitan reforzar la validación de los prototipos/ productos Web desarrollados.

3.2. Resultados Obtenidos

En la línea de investigación orientada a Usuarios Adultos Mayores, se trabajó fuertemente con este grupo de usuarios de interés en la identificación de necesidades y preferencias que mejoren el acceso e interacción con la Web [12][14]; y en particular, durante los

últimos meses de 2017, los esfuerzos estuvieron enfocados en mejorar la UX con las redes sociales y desde dispositivos móviles del tipo *tablets* [16][17]. Esta línea de investigación, contó con el soporte que ofrece el dictado anual de cursos y talleres de extensión (a cargo de integrantes de GIFIS), destinados a la población Adulta Mayor de la región. Las líneas LI.1 Adultos Mayores, LI.2 Experiencia de Usuario (UX) y LI.3 Web Móvil, se enriquecieron mutuamente con el trabajo desarrollado, ya que se revisaron y aplicaron herramientas conceptuales y prácticas en el diseño y rediseño de productos Web que permitieron desde identificar una estrategia completa de contenidos para mejorar la forma de comunicación de un sitio con sus visitantes hasta definir los componentes necesarios para definir un enfoque de diseño UX que contemple las necesidades y preferencias de grupos de usuarios de interés desde dispositivos móviles [10][13][16][17].

Otros resultados a destacar consisten en la generación de proyectos de extensión y vinculación para llegar a la comunidad regional:

- Proyecto de Extensión y Vinculación (PEyV) “Mejoras en la Accesibilidad Web de un Portal Universitario: Evaluación con Usuarios Ciegos”. En este PEyV participan alumnos de grado y postgrado.
- Proyecto de Extensión “Abuelos 2.0: Taller de Redes Sociales para Adultos Mayores”

Los integrantes del GIFIS también participaron con ponencias en los siguientes eventos locales, nacionales e internacionales:

- XIV Semana de la Ciencia, la Tecnología y el Arte Científico¹: charla “Principios Básicos de Internet para Adultos Mayores” y taller “Evaluación de la Experiencia de Usuarios en la Web”, ambas actividades abiertas a la comunidad en general.
- XVIII y XIX WICC, se presentaron los artículos [9][15] y se realizaron las ponencias con sus respectivos posters.
- 4º EIPA, se presentaron los artículos [11][12][13] y se realizaron las ponencias con sus respectivos posters.

- INCISCOS 2016, en esta conferencia internacional se presentó el artículo [10] y se viajó a Ecuador a exponerlo.
- CONAIISI 2017, en este congreso nacional se presentó el artículo [16] y se viajó a exponerlo.

4. FORMACIÓN DE RECURSOS HUMANOS

Es importante señalar que GIFIS está enfocado en brindar el marco adecuado para que cada integrante alcance sus objetivos particulares de investigación y formación al contribuir con el objetivo del PI en curso.

El PI cuenta con 4 integrantes alumnas de posgrado, las cuales están cursando la Maestría en Informática y Sistemas (MIS-UNPA). Durante 2017, estas alumnas tuvieron los siguientes avances:

- 1 alumna estuvo trabajando en la escritura de su Tesis de la MIS-UNPA.
- 1 alumna terminó su estudio preliminar del estado-del-arte con un artículo ICT-UNPA aprobado; también completo una beca de “Iniciación a la Investigación para Alumnos de Postgrado UNPA”, la cual tuvo como producción la entrega a revisión de otro artículo ICT-UNPA; además, presentó un artículo con su propuesta de Tesis en el CONAIISI 2017 y viajó a exponer este trabajo de investigación; y durante 2018, se abocará a escribir su tesis de la MIS-UNPA.
- 2 alumnas comenzaron sus Tutorías de la MIS-UNPA.

También el PI cuenta con 2 alumnos de grado, los cuales, durante 2017, desarrollaron las siguientes becas, destinadas a sus respectivas formaciones como investigadores:

- 1 alumno de la carrera Ingeniería en Sistemas UNPA estuvo trabajando en su beca de “Estímulo a las Vocaciones Científicas (EVC-CIN).”
- 1 alumno de la carrera Analista de Sistemas UNPA estuvo trabajando en su beca de “Iniciación a la Investigación para Alumnos de Grado y Pregrado UNPA.”

5. AGRADECIMIENTOS

A la UNPA por el soporte al Proyecto de Investigación (PI) N°: 29/B194, Período: 2016-2018, Denominado: “*Un Enfoque Integrador*”

¹< <http://www.semanadelaciencia.mincyt.gob.ar/>>

para Diseñar y Evaluar Interfaces de Usuario Web”.

6. REFERENCIAS

- [1] Brajnik, G. Barrier Walkthrough: Heuristic Evaluation Guided by Accessibility Barriers. <<http://users.dimi.uniud.it/~giorgio.brajnik/projects/bw/bw.html>>
- [2] Hartson, R., Pyla, P. S. The UX Book: Process and Guidelines for Ensuring a Quality User Experience. Morgan Kaufmann; 1st Edition, March 2012.
- [3] Vieritz, H., Schilberg, D., Jeschke, S. Merging Web Accessibility and Usability by Patterns, Computers Helping People with Special Needs. LNCS Vol. 6179, 2010.
- [4] UWEN (Unified Web Evaluation Methodology) Metodología de Evaluación Europea de la Accesibilidad <<http://www.wabcluster.org/uwem1/>>
- [5] Redish, J. G. Letting Go of the Words: Writing Web Content that Works. Second edition. Ed: Morgan Kaufmann; 2012.
- [6] Halvorson, K.; Rach, M. Content Strategy for the Web. Ed: New Riders; 2012.
- [7] Gothelf, J. Lean UX: Applying Lean Principles to Improve User Experience. Ed: O’Reilly; 2013. ISBN: 978-1-449-31165-0.
- [8] Levin, M. Designing Multi-Device Experiences: An Ecosystem Approach to User Experiences Across Devices. Ed: O’Reilly; 2014. ISBN: 978-1-449-34038-4.
- [9] Martín, A., Gaetán, G., Saldaño, V., Miranda, G., Sosa, H., Pires, A., Nichele, E. Evaluaciones de Accesibilidad y Usabilidad en la WWW: Propuestas para Mejorar la Experiencia del Usuario. WICC 2016; ISBN: 978-950-698-377-2; Entre Ríos.
- [10] Gaetán, G., Martín, A., Saldaño, V. Content Analysis: A Strategic Foundation to Improve the User Experience of a University Website. INCISCOS 2016; ISBN: 978-9978-389-32-4; Ecuador.
- [11] Martín, A., Gaetán, G., Saldaño, V., Cardozo, C., Miranda, G., Sosa, H. Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web. 4^o EIPA 2016; Caleta Olivia.
- [12] Cardozo, C., Saldaño, V., Martín, A., Gaetán, G. Los Adultos Mayores y la Utilización de Redes Sociales en Dispositivos Móviles. 4^o EIPA 2016; Caleta Olivia.
- [13] Miranda, G., Gaetán, G., Martín, A., Saldaño, V. Un Enfoque de Desarrollo Basado en la Experiencia de Usuario (UX) para favorecer la Accesibilidad Web. 4^o EIPA 2016; Caleta Olivia.
- [14] Cardozo, C., Martín, A., Saldaño, V. Los Adultos Mayores y las Redes Sociales: Analizando Experiencias para Mejorar la Interacción. ICT-UNPA 2017, Vol.9, N°2, P 1-29. ISSN: 1852-4516.
- [15] Martín, A., Gaetán, G., Saldaño, V., Pires, A., Miranda, G., Villagra, S., Carrizo, A., Cardozo, C., Sosa, H. Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web. WICC 2017; ISBN: 978-987-42-5143-5; Buenos Aires.
- [16] Cardozo, C., Martín, A., Saldaño, V. Abuelos del Siglo XXI: Una Propuesta para Mejorar la Experiencia en el Uso de Redes Sociales desde Dispositivos Móviles. 5to. CONAISI 2017, 2-3 Noviembre; Santa Fé; Argentina; pp. 478-487; ISSN: 2347-0372.
- [17] Cardozo, C., Martín, A., Saldaño, V. Diseño de Interfaces para Dispositivos Móviles. Una Mejora de la Experiencia de Usuarios Adultos Mayores en Redes Sociales. ICT-UNPA 2018, Beca de Investigación para Alumnos de Postgrado 2017 (trabajo en revisión de pares).

Ingeniería de software para sistemas críticos ferroviarios

Irrazábal, Emanuel; Bernal, Rubén; Pinto Luft, Cristian; Sambrana, Iván

Departamento de Informática. Facultad de Ciencias Exactas y Naturales
y Agrimensura. Universidad Nacional del Nordeste
{eirrazabal}@exa.unne.edu.ar, rbernal73@gmail.com, cristianpl777@gmail.com,
sambranaivan@gmail.com

Resumen

Esta línea de investigación aborda temas de ingeniería del software en sistemas tradicionales y su traslado a sistemas críticos con aplicación en el sistema ferroviario argentino. El propósito de esta línea de trabajo es desarrollar los sistemas de gestión para el desarrollo de proyectos software ferroviarios de acuerdo con la normativa europea EN 50128 y la construcción del ecosistema de herramientas que lo instrumenten. Esta normativa describe un conjunto de buenas prácticas y técnicas opcionales que serán tenidas en cuenta a lo largo del desarrollo software. Debido a ello es necesario analizar las posibles implementaciones de estas buenas prácticas para que sea conforme la normativa. Finalmente, la línea de trabajo plantea la posibilidad de trasladar los resultados a proyectos específicos en conjunto con la Autoridad Ferroviaria Nacional.

Palabras clave: Ingeniería de software, sistemas críticos, EN 50128, desarrollo de procesos.

Contexto

La línea de Investigación y Desarrollo presentada en este trabajo corresponde al proyecto PI-F17-2017 “Análisis e implementación de tecnologías emergentes en sistemas computacionales de aplicación regional.”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2018-2021.

Asimismo, parte de la línea de investigación es realizada es desarrollada en el marco de la tesis del maestrando Cristian Pinto Luft perteneciente a la Maestría de Tecnologías de la Información Rs. 764/14 CS UNNE.

Introducción

El sistema público ferroviario argentino se encuentra centralizado, y aunque se percibe como poco importante constituye un eslabón fundamental para la industria. En Argentina cada día tres millones de personas viajan en tren o subte y el 10% del PBI se moviliza por ferrocarril [1]. Sin embargo, todos los sistemas electrónicos para la seguridad vial de trenes y subtes son importados y muy caros. Por ejemplo, un sistema de barrera automático cuesta hasta 200.000 dólares y un sistema de control de velocidad más de 100.000 dólares. Así, en muchos trenes, no hay sistemas de seguridad para pasajeros, conductores, peatones y automovilistas y en otros, se siguen usando tecnologías de hace más de 50 años, que en los países con alto desarrollo tecnológico han sido reemplazadas hace mucho tiempo [2]. Esta situación ha favorecido que ocurran terribles accidentes [3] y ha urgido al Estado a adquirir en el exterior trenes y sistemas de seguridad ferroviaria, lo que implica enormes gastos en dólares y depender de tecnología extranjera [4][5][6]. Pero en la mayoría de los casos los accidentes se podrían haber evitado mediante el uso de sistemas electrónicos apropiados, que hoy en día son habituales en países con alto desarrollo tecnológico. Sin embargo, como se mencionó

anteriormente, en la actualidad estos sistemas no se desarrollan en la Argentina.

Existen, sin embargo proyectos de investigación y de extensión que se encuentran actualmente trabajando en ello. Un ejemplo de ello es el Proyecto Desarrollo de Estratégico UBA N°23 "Controlador electrónico para barreras automáticas ferroviarias con nivel de integridad de seguridad certificable hasta SIL4" desarrollado por el Dr. Ariel Lutenberg, director del Programa CIAA. El objetivo de este proyecto ha sido desarrollar un prototipo de Monitor de Barrera ferroviaria construido a partir de normas internacionales y componentes electrónicos programables, en este caso la Computadora Industrial Abierta Argentina.

Los sistemas ferroviarios son complejos, compuestos por distintos componentes software, hardware y humanos, que interactúan con su entorno de maneras muy variadas. Un fallo en uno de estos componentes o subsistemas puede llegar a tener asociados distintos niveles de peligros, pudiendo causar pérdidas financieras, daño al equipamiento, daños ambientales, lesiones a personas o en los peores casos pérdidas de vidas humanas. Por estos motivos dichos sistemas se encuentran regulados con distintas leyes y normativas cuyo fin es preservar los recursos anteriormente mencionados [7]. Algunos de los principales organismos que regulan esta actividad son el Comité Européen de Normalisation Electrotechnique (CENELEC) en Europa o la International Electrotechnical Commission (IEC) en América.

Una de las características más importantes de los sistemas que estas normas intentan reforzar durante todo su ciclo de vida son las de fiabilidad, disponibilidad, mantenibilidad y seguridad (RAMS por sus siglas en inglés).

Las principales normas propuestas por el CENELEC orientadas a la resolución de la problemática explicada anteriormente son las siguientes:

- EN 50126 [8]: Aplicaciones ferroviarias. La especificación y demostración de Fiabilidad, Disponibilidad, Mantenibilidad y Seguridad (RAMS). Esta norma se orienta principalmente al cumplimiento de las características RAMS del sistema en general.
- EN 50128 [9]: Aplicaciones ferroviarias. Sistemas de comunicación, señalización y procesamiento. Software para sistemas de control y protección del ferrocarril. Esta norma se centra principalmente en la calidad de los aspectos software de los sistemas de ferrocarriles.
- EN 50129 [10]: Aplicaciones ferroviarias. Sistemas de comunicación, señalización y procesamiento. Sistemas electrónicos relacionados con la seguridad para la señalización. Esta norma se centra principalmente en los aspectos de calidad del hardware de los sistemas de ferrocarriles.

Una de las características principales de los sistemas críticos es la seguridad de la que estos deben estar dotados por naturaleza, debido a las consecuencias que pueden provocar sus fallos. Para dotarlos de seguridad, una de las metodologías utilizadas en su diseño es el aseguramiento de los mismos desde su concepción, es decir, desde el análisis y definición de sus requerimientos. Para esto se utilizan enfoques que integran las disciplinas de ingeniería de requerimientos con la ingeniería de seguridad, lo cual está comprobado que aumenta significativamente la seguridad del sistema en sí [11].

En la actualidad, grandes organizaciones como la NASA [12], Ansaldo Signal o Siemens Rail Transportation [13] utilizan una combinación de metodologías y formas de trabajo provenientes de distintos campos del conocimiento para lograr dicha vinculación, y de esta manera mejorar la calidad y seguridad de los sistemas críticos que desarrollan, dedicando tiempo, recursos y esfuerzo a esta tarea. Esta línea de trabajo utilizará principalmente los enfoques

propuestos por las normas UNE-EN 50126:2005 y UNE-EN 50128:2012, gestionando las políticas RAMS vinculadas a los requerimientos de los subsistemas software que componen a los sistemas ferroviarios, y haciendo hincapié en los aspectos de seguridad de las mismas.

Para ello es indispensable el desarrollo de procedimientos y un sistema de gestión de calidad conforme la normativa UNE-EN 50128. Esta normativa detalla las buenas prácticas de las diferentes fases en el desarrollo de los sistemas software ferroviarios en las aplicaciones de señalización y control.

Centrándose en la gestión de la seguridad de los requisitos software, existen una amplia variedad de técnicas con las que se intenta, en distintos niveles y bajo diferentes enfoques, dotarlos de dicho atributo de calidad. Entre ellas se pueden encontrar Preliminary Hazard Analysis (PHA) [14], Software Failure Modes and Effect Analysis (SFMEA) [15], Software Effect and Criticality Analysis (SFMECA) [15], Software Failure Tree Analysis (SFTA) [15], Software Common Cause and Failure Analysis (SCCFA) [15], Hardware-Software Integration Analysis (HSIA) [15], Software Subsystem Hazard Analysis (SSHA) [12], Deductive Cause Consequence Analysis (DCCA) [16], entre otras.

Asimismo, existen diferentes herramientas que intentan dar soporte a dicha gestión, cubriendo determinados aspectos relacionados al modelado, control y chequeo de cuestiones relativas a la seguridad, como ser por ejemplo SCADE [17], Matlab [18], CodeCheck [18], StackAnalyzer [19], aiT WCET Analyzer [19], Astrée [19], OVADO [20], Atelier B [21], Rodin [21], Verasis [21], Eclipse Modeling Framework [22], MOFScript [22], PolySpace [23], entre otras. Además de utilizarse herramientas para realizar estos procesos, también se desarrollan y utilizan diversos modelos y meta-modelos de procesos, tratando de abarcar todas las

sub áreas que componen al tema en cuestión [16][24][25][26].

Otras herramientas muy utilizadas en este campo de investigación y desarrollo son los métodos formales de modelado, como ser: Formal Failure Model [16], redes de Petri [27], método formal B (lenguaje B) [20][21] o Abstract Interpretation [19][23] entre otros. Mediante los mismos se pretende disminuir la ambigüedad de las distintas fases relacionadas a la gestión de seguridad de los requerimientos software, partiendo desde su análisis, pasando por el diseño y llegando a la verificación formal de los productos a desarrollar.

El propósito de esta línea de trabajo es, por tanto, desarrollar los sistemas de gestión y la construcción del ecosistema de herramientas que lo instrumenten, así como su verificación y validación en ensayos junto con la Autoridad Ferroviaria Nacional, valiéndose de las investigaciones realizadas en el marco del proyecto.

Líneas de investigación y desarrollo

En la línea de Ingeniería de Software para Sistemas Críticos se propone:

- Estudiar las normativas de sistemas críticos ferroviarios para el desarrollo de firmware y software certificables.
- Desarrollar el conjunto de procedimientos para la gestión de requerimientos en entornos críticos y de seguridad funcional.
- Desarrollar un sistema de gestión de calidad adaptados a entidades que construyen sistemas embebidos y sistemas críticos de acuerdo con la norma internacional EN 50128 e ISO 9001.
- Validar los procedimientos construyendo prototipos de sistemas a ser utilizados por la Autoridad Ferroviaria Nacional.

Resultados obtenidos

El grupo de investigación es de reciente formación, por lo cual los resultados son preliminares y, en parte, se enumeran antecedentes llevados adelante en el marco de otros grupos de trabajo. A continuación se indican:

En la línea de Ingeniería de Software para Sistemas Críticos:

- Se está trabajando en el desarrollo de un Ecosistema de Calidad de Software para Sistemas Críticos a partir de las herramientas Jenkins, SONAR, Eclipse Process Framework (EPF), Redmine y Testlink. Estas herramientas sirven como soporte para la gestión colaborativa de los proyectos, la descripción de los procedimientos y el análisis del código fuente [28].
- Se está trabajando en la construcción de los procedimientos de desarrollo software de acuerdo con la norma UNE-EN 50128 soportado por EPF [29].

Formación de recursos humanos

En el Grupo de Investigación en Innovación en Software y Sistemas Computacionales (GISSC) están involucrados 4 docentes investigadores, 1 becario de investigación de pregrado, 1 tesista de doctorado y 3 tesistas de maestría. Cinco alumnos de la carrera están realizando sus proyectos finales vinculado a estos temas.

Para el caso de esta línea de investigación se encuentran trabajando dos docentes investigadores, un tesista de maestría y dos tesistas de pregrado.

Referencias

- [1] Sitio web con estadísticas CNRT: <https://www.cnrt.gob.ar/content/estadisticas>, Visitado: 19/02/2018.
- [2] Ingeniería y gestión del mantenimiento en el sector ferroviario (Spanish Edition) Paperback – 2010, Arques Paton José Luis.
- [3] Sitio web con ejemplos de accidentes ferroviarios argentinos:: https://es.wikipedia.org/wiki/Categor%C3%ADa:Accidentes_ferrovianos_en_Argentina. Visitado: 19/02/2018,
- [4] Sitio web com ejemplo de licitación: https://www.clarin.com/ciudades/tren-es-china-compra-licitacion-reestatizacion_0_rJduN7cwXe.html. Visitado:19/02/2018.
- [5] https://www.clarin.com/ieco/china-trenes_de_carga-randazzo-inversiones_0_rJygK8mKP7l.html
- [6] <https://www.argentina.gob.ar/noticias/japon-comenzara-fabricar-la-tecnologia-para-el-frenado-automatico-de-trenes>
- [7] J. L. Boulanger, “CENELEC 50128 and IEC 62279 Standards”, Control, Systems and Industrial Engineering Series, John Wiley & Sons, Inc., 2015, p. 13.
- [8] EN 50126. Railway applications - The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS). 2005.
- [9] EN 50128. Railway applications - Communication, signalling and processing systems - Software for railway control and protection systems. 2011.
- [10] EN 50129. Railway applications - Communication, signalling and processing systems - Safety related electronic systems for signaling. 2005.
- [11] J. Vilela, J. Castro, L. E. G. Martins, T. Gorschek, “Integration between Requirements Engineering and Safety Analysis: A Systematic Literature Review”, The Journal of Systems & Software, Vol. 125, Pp. 68-92, Marzo, 2017.
- [12] NASA Software Safety Guidebook. NASA Technical Standard. NASA-GB- 8719.13. Marzo, 2004. [15] Ansaldo STS. Website, última visita 11/07/2017. <http://www.ansaldo-sts.com/en/index>

- [13] [Siemens Rail Transportation. Website, <http://w3.siemens.com/mcms/industrial-controls/en/railway/Pages/overview.aspx> Visitado 19/02/2018
- [14] J. Kotti, S. Panchumarthy, "The Quantitative Safety Assessment and Evaluation for Safety-Critical Computer Systems", SIGSOFT Softw. Eng. Notes, Vol. 41, pp. 1-8, Enero, 2016.
- [15] R. Pietrantuono, S. Russo, "Introduction to Safety Critical Systems", Innovative Technologies for Dependable OTS-Based Critical Systems, pp. 17-27, Enero, 2013.
- [16] F. Ortmeier, M. Gudemann, W. Reif, "Formal Failure Models", IFAC Proceedings Volumes, Vol. 40, pp. 145-150, Junio, 2007.
- [17] M. Huhn, S. Milius, "Observations on formal safety analysis in practice", Science of Computer Programming, vol. 80, pp. 150-168, Febrero, 2014.
- [18] T. L. Johnson, H. A. Sutherland, B. Ingleston, B. H. Krogh, "Dependable Software in Railway Signalling", IFAC Proceedings Volumes, vol. 38, pp. 42-49, 2005.
- [19] D. Kästner, C. Ferdinand, "Applying Abstract Interpretation to Verify EN-50128 Software Safety Requirements", Reliability, Safety, and Security of Railway Systems. Modelling, Analysis, Verification, and Certification, LNCS, vol. 9707, pp. 191-202, Junio, 2016.
- [20] R. Abo, L. Voisin, "Formal Implementation of Data Validation for Railway Safety-Related Systems with OVADO", Software Engineering and Formal Methods, LNCS, vol. 8368, pp. 221-236, Marzo, 2014.
- [21] A. G. Russo Jr., "Formal Methods as an Improvement Tool", Industrial Deployment of System Engineering Methods, pp. 81-95, 2013.
- [22] A. Svendsen, G. K. Olsen, J. Endresen, T. Moen, E. Carlson, K.-J. Alme, Ø. Haugen, "The Future of Train Signaling", Model Driven Engineering Languages and Systems, LNCS, vol. 5301, pp. 128-142, 2008.
- [23] A. Ferrari, D. Grasso, G. Magnani, A. Fantechi, M. Tempestini, "The Metrô Rio ATP Case Study", Formal Methods for Industrial Critical Systems, LNCS, vol. 6371, pp. 1-16, 2010.
- [24] J. L. de la Vara, A. Ruiz, K. Attwood, H. Espinoza, R. Kaur Panesar-Walawege, Á. López, I. del Río, T. Kelly, "Model-based specification of safety compliance needs for critical systems: A holistic generic metamodel", Information and Software Technology, vol. 72, pp. 16-30, Abril, 2016.
- [25] M. Huhn, H. Hungar, "8 UML for Software Safety and Certification", Model-Based Engineering of Embedded Real-Time Systems, LNCS, vol. 6100, pp. 201-237, 2010.
- [26] D. Fowler, P. Bennett, "IEC 61508 - A Suitable Basis for the Certification of Safety-Critical Transport-Infrastructure Systems", Computer Safety, Reliability and Security, LNCS, vol. 1943, pp. 250-263, 2000.
- [27] M. S. Durmus, U. Yildirim, O. Eris, M. T. Söylemez, "Safety-Critical Interlocking Software Development Process for Fixed-Block Signalization Systems", IFAC Proceedings Volumes, vol. 45, pp. 165-170, Septiembre, 2012.
- [28] Diapositivas y explicaciones del funcionamiento del Ecosistema de Calidad em www.linsse.com.ar. Visitado: 18/02/2018
- [29] Sitio web construído en EPF www.linsse.com.ar/epf. Visitado: 18/02/2018

“Desarrollo e implementación de un modelo basado en estándares para el logro y evaluación de la Usabilidad en aplicaciones informáticas”

CASTRO Marcelo, SÁNCHEZ RIVERO David, VARGAS Alejandro, ARAGÓN Fabiana, REINOSO Elizabeth, APARICIO María, FARFÁN José, CÁNDIDO Andrea, CAZÓN Liliana & ZAPANA, José

Investigación + Desarrollo en Gobierno Electrónico / Facultad de Ingeniería /
Universidad Nacional de Jujuy

Av. Italia y Av. Martiarena / S. S. de Jujuy / Provincia de Jujuy

Tel. 388-4221591

mcastro@fi.unju.edu.ar, vdsanchezrivero@fi.unju.edu.ar, lavargas@fi.unju.edu.ar,
fraragon@fi.unju.edu.ar, edrreinoso@fi.unju.edu.ar, mcaparicio@fi.unju.edu.ar,
jhfarfan@fi.unju.edu.ar, agcandido@fhyics.unju.edu.ar, ddcastro@fi.unju.edu.ar,
lbcazon@fce.unju.edu.ar, jvzapana@fi.unju.edu.ar

RESUMEN

El presente proyecto de investigación tiene por objeto diseñar, desarrollar e implementar un modelo basado en estándares internacionales relacionados con la calidad: ISO 9241, ISO/IEC 9126-1, ISO/IEC 9126-4, ISO/IEC 25010, WCAG 2.0 e ISO/IEC 25062, para el logro y evaluación de la Usabilidad en aplicaciones informáticas teniendo en cuenta el alcance, la dimensión y la relación con otros términos de calidad en las interfaces de usuario; accesibilidad, calidad en uso y experiencia de usuario(UX).

Además se estudiarán y se incorporarán para el diseño del modelo, los conceptos sobre Ingeniería de la usabilidad, Diseño centrado en el uso, patrones de diseño de interfaces de usuario y Desarrollo dirigido por modelos (MDD). En relación a la evaluación de la Usabilidad se estudiarán los distintos métodos existentes, haciendo hincapié en las recomendaciones del estándar ISO/IEC 25062:2006, además se utilizarán metodologías y herramientas relacionadas con la evaluación de la accesibilidad.

Para probar el modelo desarrollado se creará un Laboratorio de Usabilidad que dependerá del Gabinete de Herramientas de Software, ubicado en el centro de cómputos de la Facultad de Ingeniería de la UNJu. El Laboratorio de Usabilidad proveerá servicios a distintas entidades públicas y privadas que

deseen evaluar la usabilidad de sus productos informáticos.

Palabras clave: Usabilidad, Accesibilidad, Diseño Centrado en el Usuario (DCU), Experiencia de usuario(UX), Calidad en uso, Evaluación de software.

CONTEXTO

El proyecto “Desarrollo e implementación de un modelo basado en estándares para el logro y evaluación de la Usabilidad en aplicaciones informáticas” es una continuación e incorpora conceptos y experiencias de los proyectos desarrollados por nuestro equipo de investigación denominados, “Tic’s: Automatización y Estandarización del Proceso de Gobierno Electrónico”, entre los años 2007 al 2009, “Desarrollo de una herramienta para automatizar el proceso de Gobierno Electrónico”, cuyo periodo de trabajo comprendió los años 2010 al 2012 y “Desarrollo de un modelo basado en servicios digitales comunes reutilizables para Gobierno Electrónico” desde el año 2013 al 2015. Estos constituyeron la base conceptual para elaborar un modelo que permitiera, basado en estándares logro y evaluación de la Usabilidad en aplicaciones informáticas.

El proyecto se desarrolla en la Facultad de Ingeniería de la Universidad Nacional de Jujuy, posee código D-0142 y posee Categoría “A”. Además se encuentra subsidiado por la Secretaría de Ciencia y Técnica y Estudios Regionales, dependiente de la U.N.Ju.

1. INTRODUCCIÓN

En el modelo para la evaluación de la usabilidad basado en estándares que se propone en el proyecto de investigación, se analizarán las diferencias y relaciones existentes entre los conceptos de usabilidad, diseño centrado en el usuario, facilidad de uso, accesibilidad, ergonomía, experiencia de usuario, calidad en uso y usabilidad universal.

La Usabilidad significa que la gente que utilice un producto pueda realizar rápida y fácilmente sus tareas. [1]

La usabilidad realmente significa tener la certeza de que algo funciona bien, es decir, que cualquier persona con conocimientos o habilidades medias (incluso por debajo de la media) puede navegar por un sitio web, usar un avión o abrir una puerta, sin que cualquier labor que lleve a cabo conlleve su frustración. [2]

La Usabilidad universal es la posibilidad y grado con el que usuarios diversos, con tecnologías diferentes e intereses también distintos pueden acceder a la información y a los servicios. [3]

Es importante tener presente que la usabilidad no es algo simple o una propiedad que ofrezca una única cara relacionada con la interfaz de usuario ofrecida. La usabilidad tiene múltiples componentes e involucra cinco atributos: aprendizaje, eficiencia, memorización, tratamiento de errores y satisfacción. [4]

La accesibilidad Web significa que personas con algún tipo de discapacidad van a poder hacer uso de la Web. En concreto, al hablar de accesibilidad Web se está haciendo referencia a un diseño Web que va a permitir que estas personas puedan percibir, entender, navegar e interactuar e interactuar con la Web, aportando a su vez contenidos. [5]

La ergonomía es el conjunto de conocimientos de carácter multidisciplinar aplicados para la adecuación de los productos, sistemas y entornos artificiales a las necesidades, limitaciones y características de sus usuarios, optimizando la eficacia, seguridad y bienestar. [6]

La experiencia de Usuario (UX) es la respuesta a cómo se sienten los usuarios después de interactuar con un sistema. [7] [8]

[9]

La calidad en uso, es un concepto que surge desde la ingeniería del software y viene a salvar las diferencias entre ésta y la interacción persona-ordenador (HCI).

1.1. Estándares Internacionales

Para la característica de usabilidad los estándares internacionales plantean la calidad del producto software y la calidad de la interacción, es decir calidad del producto y calidad del proceso. A su vez las disciplinas que trabajan estrechamente con la usabilidad, son la Ingeniería del Software y la Interacción Hombre Máquina (HCI).

En el caso de la Ingeniería del software se preocupa por la calidad del producto software a través de la evaluación de características tales como la funcionalidad, la Fiabilidad, la eficiencia, la mantenibilidad, la portabilidad, y de la Usabilidad. La Interacción Hombre Máquina, en cambio, se preocupa por la calidad del proceso de interacción, verificando la accesibilidad, la experiencia de usuario y la usabilidad.

Los estándares internacionales trabajan con diversos factores de calidad asociados tanto al producto software como al proceso de interacción que ofrecen los mismos. En la tabla 1, se puede apreciar la relación entre los distintos estándares internacionales y los factores de calidad que consideran.

Estándar	Factores de calidad considerados
ISO9241-11	Usabilidad
ISO 9241-210	Usabilidad, UX y DCU
ISO 9241-171	Accesibilidad
ISO/IEC9126-1	Funcionalidad, eficiencia, fiabilidad, usabilidad, portabilidad y mantenimiento
ISO/IEC 9126-4	Calidad en uso
ISO/IEC 25010	Calidad de un producto software y de la interacción de dicho producto

Tabla 1 – Relación entre Estándar y Factores de calidad

El estándar ISO 9241-11 [10] establece que la usabilidad es efectividad, eficiencia y satisfacción. Efectividad en el sentido que el usuario es capaz, utilizando el producto software, de lograr sus objetivos. Eficiencia ya

que logra sus objetivos y lo hace destinando los recursos necesarios (tiempo, aprendizaje, etc.). Satisfacción del usuario en el uso del producto software tornándose placentero, confortable y útil.

En el estándar ISO 9241-210 [11] se definen los conceptos y las técnicas asociadas a Usabilidad, Experiencia de Usuario (UX) y Diseño Centrado en el Usuario (DCU).

En la norma ISO 9241-171 [12] se identifican un conjunto de requisitos relacionados con el logro de la accesibilidad. Brinda información consistente con documentos de la W3C relacionados con accesibilidad en la Web (WCAG 2.0.)

La ISO/IEC 9126-1 [13] establece las características, y sub-características que se deben tener en cuenta para evaluación de la calidad interna y externa de un producto software.

El estándar ISO/IEC 9126-4 [14] introduce el concepto de calidad en uso, relacionado con la calidad de la interacción. Además incorpora métricas para su estimación y anexos relacionados con la elaboración de informes sobre la calidad en uso/usabilidad de un producto software.

El ISO/IEC 25010 [15] sustituye a la ISO/IEC 9126-1, y sigue considerando la usabilidad como producto software. Mantiene algunos criterios tradicionales relacionados con usabilidad, pero agrega criterios adicionales como el riesgo y la adecuación al contexto.

1.2. Evaluación de la Usabilidad

Existen técnicas que dan soporte a las actividades de evaluación de interfaces de usuario, por ejemplo las descritas por el estándar ISO 9241-210 [11], éstas incluyen las actividades de planificar el proceso de diseño, entender y especificar el contexto de uso, identificar y especificar los requisitos del usuario, diseñar e implementar la interface de usuario.

Algunos métodos de evaluación de la usabilidad y soporte al DCU [16] [17] se pueden observar en la tabla 4.

Nombre del método	Breve descripción
Observación de usuarios	Recogerde manera sistemática información sobre el comportamiento y las prestaciones de los usuarios en un contexto específico mientras realizan su actividad.
Medición de prestaciones	Colección de medidas con la intención de conocerel impacto de los aspectos de usabilidad.
Análisis de incidentes críticos	Recolecciónde eventos específicos (positivos y negativos).
Cuestionarios	Métodos de evaluación indirectaque consiguen opiniones de los usuarios sobre la interfaz de usuario utilizando cuestionarios predefinidos.
Entrevistas	Similar a los cuestionarios con mayor flexibilidad y utilizando una interacción cara a cara.
Thinkingaloud	Involucrar a usuariosy pedirles que verbalicen sus opiniones, creencias, dudas, etc. mientras interactúan con el sistema.
Evaluación y diseño colaborativo	Métodos que permiten a diferentes tipos de usuarios colaborar enla evaluación y diseño de sistemas.
Métodos creativos	Métodos que persiguenla elicitaciónde nuevos productos con nuevas características.
Métodos basados en documentos	Examen de documentos existentes por especialistas para formarseun juicio profesional del sistema.
Desarrollos basados en modelos	Utilización de representacionesabstractas del producto evaluado para permitir la predicción de las prestaciones de los usuarios.
Evaluación experta	Evaluación basada en el conocimiento de expertos.
Evaluación automática	Evaluación dirigida por algoritmos centrados en criterios de usabilidad y conocimiento ergonómico con los que diagnosticar deficiencias.

Tabla 4 - Métodos de evaluación de la usabilidad y soporte al DCU

Para el desarrollo del modelo de evaluación a generar se propone trabajar con los estándares descriptos precedentemente y en especial con la norma ISO/IEC 25062 [18] que define el CIF (Common Industry Format) como estándar a utilizar para confeccionar informes de pruebas de usabilidad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El equipo de trabajo se encuentra desarrollando actividades de investigación y desarrollo, fundamentalmente en el área de la Ingeniería de software y específicamente en la formalización del proceso de gobierno electrónico, Ingeniería Web, Calidad del Software y Arquitectura dirigida por Modelos (MDA) a partir de los proyectos: “TIC’s: automatización y estandarización del proceso de Gobierno Electrónico”, “Diseño y Desarrollo de una herramienta para automatizar el proceso de Gobierno Electrónico”, “Desarrollo de un modelo basado en MDA para gobierno electrónico” y “Desarrollo de un modelo basado en servicios digitales comunes reutilizables para Gobierno Electrónico”.

En el nuevo desafío que presenta el proyecto “Desarrollo e implementación de un modelo basado en estándares para el logro y evaluación de la Usabilidad en aplicaciones informáticas”, se trabaja sobre ejes disciplinares relacionados a la utilización de estándares internacionales en aspectos relacionados a la usabilidad.

En la actualidad la demanda de aplicaciones informáticas requiere el desarrollo de interfaces de usuario de calidad. Para ello los diseñadores y los desarrolladores deben considerar múltiples requisitos, pero fundamentalmente la usabilidad del producto. En consecuencia se hace necesario abordarla como un proceso ingenieril para lograr los mejores resultados y poder evaluarla.

A través de estudios preliminares realizados por el grupo de investigación en los proyectos ejecutados desde al año 2007 y descriptos anteriormente, se detectó que las interfaces de usuario de diversas aplicaciones informáticas, particularmente las desarrolladas en ambientes Web tanto en el sector público como en el empresarial, no fueron concebidas y evaluadas, teniendo en cuenta estándares internacionales relacionados a la usabilidad del producto software.

Sumado a lo expuesto precedentemente, se puede decir que organizaciones

gubernamentales y no gubernamentales relacionadas a brindar ayuda y servicios a personas con capacidades diferentes, no cuentan con instalaciones especiales y procesos para realizar pruebas de accesibilidad a aplicaciones informáticas, que abarcan a interfaces de usuario de software basados en la Web, como a aplicaciones del tipo HCI.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se espera que el modelo a desarrollar en el presente proyecto de investigación como el laboratorio a montar en el ámbito de la Facultad de Ingeniería de la UNJu, tendrán un rol importante en la Provincia y en la región, brindando servicios a una gran parte de la sociedad, que va desde empresas dedicadas al desarrollo del software, pasando por compañías que desarrollan sus propias aplicaciones; hasta organismos públicos que desarrollan software para brindar servicios al ciudadano e instituciones no gubernamentales que permiten el acceso a las nuevas tecnologías de la información y comunicación a personas con capacidades diferentes, tales como el Programa para personas con discapacidad de la Universidad Nacional de Jujuy (UnjuProDis), el Instituto Helen Keller, creado para la atención de niños sordos y el Centro de Rehabilitación Integral de la Asociación de Protección al Parálítico Cerebral (Appace).

Por lo expuesto, el proyecto de investigación aspira satisfacer demandas relacionadas con aspectos técnicos informáticos, como la usabilidad y la accesibilidad de un producto software, pero también a aspectos sociales brindando servicios a instituciones que trabajan con personas con capacidades diferentes.

Los resultados obtenidos hasta el momento, se encuentran plasmados en los trabajos presentados en diversos eventos científicos entre los que podemos citar “Evaluando la calidad de uso en una aplicación web de gobierno electrónico”, “Los sistemas informáticos y la innovación tecnológica” y “Evaluando la usabilidad de Galaxy Conqueror

con ISO 25062”. Los citados trabajos fueron desarrollados en el año 2016.

En el año 2017 se presentaron los siguientes trabajos “Evaluación de la Calidad de Uso en una Aplicación de la Administración Federal de Ingresos Públicos”, “La Sociedad de la Información” y finalmente “Propuesta de cambio arquitectónico Modelo-Vista-Controlador en un organismo municipal”.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto de investigación se encuentra conformado por diez docentes investigadores, los cuales se encuentran categorizados. Además cuenta con cuatro alumnos que están siendo iniciados en tareas de investigación y actividades de desarrollo.

Por otra parte el desarrollo de las tareas de investigación, ha generado en los años 2012, tres anteproyectos de tesis en la Maestría en Ingeniería de Software, de la Universidad Nacional de San Luis pertenecientes a integrantes del equipo de trabajo.

Además se realizó la presentación y aprobación de una propuesta técnica de Tesis Doctoral para cursar el Doctorado en Ciencias Informáticas en la Facultad de Informática de la Universidad Nacional de la Plata, que se encuentra en etapa de elaboración final de la tesis doctoral.

Simultáneamente algunos miembros del equipo de trabajo, se encuentran dirigiendo cuatro proyectos finales de carrera de grado, pertenecientes a alumnos de Ingeniería Informática, de la Facultad de Ingeniería de la Universidad Nacional de Jujuy.

5. BIBLIOGRAFÍA

[1] Dumas, Joseph S. and Janice C. Redish, A Practical Guide to Usability Testing, Rev. Edition, Intellect Books, 1999.

[2] Krug, Steve, Don't Make Me Think: A Common Sense Approach to the Web, 2nd Edition, New Riders Publishing, 2005.

[3] Shneiderman, Ben, Universal Usability, ACM 43(5), Pag. 84-91, 2000.

[4] Nielsen, Jakob, The Usability Engineering Life Cycle, IEEE Computer 25(3), Pag. 12-22, 1992.

[5] Accesibilidad, WorldWide Web Consortium, W3C. <http://www.w3c.es>

[6] ISO/TR16982, Ergonomics of human-system interaction -Usability methods supporting human-centred design, 2002.

[7] Norman, Donald, The Design of Everyday Things, Basic Books, 2002.

[8] Chong Law Effie Lai, Roto, Virpi, Vermeeren Arnold, Kort Joke, Hassenzahl Marc, Towards a shared definition of user experience. CHI Extended Abstracts, Pag. 2395-2398, 2008.

[9] User experience design, http://semanticstudios.com/user_experience_design/

[10] ISO 9241-11, Ergonomic requirements for office work with visual display terminals (VDTs), Part 11: Guidance on usability, 1998.

[11] ISO 9241-210, Ergonomics of human-system interaction, Part 210: Human-centred design for interactive systems, 2010.

[12] ISO/DIS 9241-171, Ergonomics of human-system interaction, Guidance on software accessibility, 2006.

[13] ISO/IEC 9126-1, Software engineering - Product quality -Part 1: Quality model, 2001.

[14] ISO/IEC TR 9126-4, Software engineering -Product quality -Part 4: Quality in use metrics, 2004.

[15] ISO/IEC 25010, Systems and software engineering -Systems and software Quality - Requirements and Evaluation (SQuaRE) - System and software quality models, 2010.

[16] Métodos de evaluación de la usabilidad. <http://www.usabilitynet.org>.

[17] MAUSE, Towards the MAuration of Information Technology USability Evaluation, [http://www.cost294.org\(digital library\)](http://www.cost294.org(digital_library)).

[18] ISO/IEC 25062, Software engineering - Software product Quality Requirements and Evaluation (SQuaRE) -Common Industry Format (CIF) for usability test reports, 2006.

Aplicación Reflexiva de un Proceso de Requisitos

Gladys N. Kaplan¹, Renata S. Guatelli¹, Jorge H. Doorn^{1,2}, Andrea F. Vera¹,
María L. Pepe¹

¹Departamento de Ingeniería e Investigaciones Tecnológicas, Universidad Nacional de la Matanza

²Escuela de Informática, Universidad Nacional del Oeste

gkaplan@unlam.edu.ar, jdoorn@exa.unicen.edu.ar,
{renata.guatelli, andreafabianavera, laurapepe}@gmail.com

RESUMEN

En informática es una práctica habitual aplicar en forma reflexiva los métodos y estrategias que están siendo desarrollados al propio proceso de desarrollo. Ya en las etapas fundacionales de esta disciplina se hizo presente la reflexividad, por ejemplo en las primeras versiones del lenguaje LISP. Sin embargo en la Ingeniería de Requisitos hay pocos ejemplos del uso de esta estrategia. El presente proyecto se enmarca en un proceso de requisitos al cual se le han realizado mejoras en forma continua, pese a lo cual el mismo es aún susceptible de ser perfeccionado. La propuesta consiste en la aplicación en forma reflexiva de los modelos del proceso al proceso mismo de construcción de esos modelos. Se ha concentrado la atención en buscar puntos y mecanismos de creación de rastros en forma automática, en definir los requisitos que debe atender una herramienta que automatice parcialmente el Proceso de Requisitos y en buscar debilidades en el mismo. Se planifica comparar estos requisitos con los servicios provistos por herramientas construidas que no han seguido este proceso.

Palabras clave: Ingeniería de Software, Ingeniería de Requisitos, Proceso de Requisitos, Lenguaje Natural.

CONTEXTO

La línea de investigación que se presenta es parte de los proyectos de investigación “Reflexividad como herramienta en la Ingeniería de Requisitos” de la Universidad

Nacional de La Matanza (UNLaM) y “Tratamiento de los Factores Situacionales y la Completitud en la Ingeniería de Requisitos” de la Universidad Nacional del Oeste (UNO).

1. INTRODUCCIÓN

Tanto en la gestión del proceso de requisitos como en la gestión de los requisitos en sí mismos, la construcción y uso de técnicas de rastreabilidad es un aspecto no atendido o pobremente abordado tanto en los estudios académicos como en la práctica cotidiana. En gran parte esto es debido a la poca disponibilidad de herramientas automáticas o semiautomáticas para identificar y mantener trazas [4] [5] [6] [7].

La recolección y mantenimiento de la información de rastreo es de muy alto costo. Por lo tanto, se deben tener políticas que indiquen qué tipo de rastreos se realizarán y cómo se mantendrá dicha información.

Por otra parte, las herramientas desarrolladas para la estrategia utilizada como base en el presente proyecto, han sido construidas utilizando un proceso de desarrollo ad-hoc, que ha desconocido parcialmente los principios del proceso al que aspiran a servir. Esta es una realidad que brinda un espacio de comparación muy importante ya que permite detectar las potenciales desviaciones entre los requisitos detectados mediante el proceso y los servicios que prestan los sistemas construidos.

Finalmente, las mejoras introducidas al proceso de requisitos, a lo largo del tiempo, parecen estar acercándose asintóticamente a

un valor límite que no ofrece garantías. Posiblemente esta evolución se deba a una falta de crítica acerca de los mecanismos básicos involucrados. Una de las formas que permite visualizar esta falta de crítica consiste en releer las heurísticas de construcción de los diferentes modelos del proceso. En casi todas ellas se puede notar que las mismas son más cercanas a una descripción de los componentes del modelo que a una guía meditada que ayude a afrontar el trabajo en forma eficaz.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A lo largo de dos décadas en varios proyectos de investigación, se ha desarrollado una estrategia de Ingeniería de Requisitos basada en modelos en lenguaje natural [8], que ha sido difundida en cursos de grado y posgrado, y puesta en práctica en diversos proyectos de software de mediana a gran envergadura en la industria. Aun cuando esta estrategia puede considerarse suficientemente madura, todavía es concebible que la misma pueda ser mejorada en varios aspectos.

Por un lado los costos en un proyecto de software originados por una inapropiada gestión de requisitos, especialmente en lo que se refiere a la inadecuada atención de los cambios en los mismos, son de una importancia grande o muy grande dependiendo del grado de volatilidad de los mismos.

Casi sin excepción, un tratamiento inadecuado de los cambios en los requisitos deviene de la carencia de registros que indiquen en forma precisa qué aspecto del proceso del negocio originó un determinado requisito y como consecuencia qué requisitos relacionados serán también afectados.

Debe notarse que los cambios en los requisitos no surgen como referencias explícitas y claras a determinadas funcionalidades del sistema de software, sino como problemáticas del dominio del problema usualmente relacionadas en forma poco clara con los requisitos del sistema en desarrollo.

Otro aspecto relevante es que en muchos

casos se consideran cambios en los requisitos a meros errores de interpretación u omisiones en las primeras fases del proyecto.

Por otro lado, la imposibilidad efectiva de crear modelos completos y el fantasma de la existencia de omisiones graves está siempre presente en todo proceso de requisitos. En estudios previos se ha comprobado la importancia de las omisiones, al extremo de carecerse de estimaciones de la completitud lograda en el proceso [9] [10] [11] [12].

Es muy tentador y numerosos autores reportan que incrementa la probabilidad de éxito de un proyecto de software, el estudio del dominio del problema y luego que los requisitos de ese sistema se plasmen en descripciones o modelos creados en Lenguaje Natural [13]. Las mejoras registradas en la calidad de los requisitos tienen su fundamento en que las representaciones en lenguaje natural, tales como glosarios, casos de uso y escenarios, promueven la comunicación entre todos los involucrados en un proyecto de desarrollo de software y facilitan la validación de los requisitos elaborados. Una revisión hecha por Rolland et al. [14] muestra que, de doce enfoques propuestos en la literatura en el ámbito de la Ingeniería de Requisitos, todos ellos usan una notación de texto para describir escenarios, que en algunos casos se combinan con otros medios, tales como gráficos o imágenes. En un estudio relativamente reciente sobre la práctica en Ingeniería de Requisitos [15], se concluyó que el 51% de las organizaciones (sobre un total de 194) usa representaciones informales (por ejemplo, el lenguaje natural) y el 27% semi-formales, quedando los modelos formales relegados a un uso de apenas el 7% de las organizaciones.

Este proyecto está enmarcado en una estrategia en Ingeniería de Requisitos orientada al cliente, pues efectivamente se basa en la construcción de modelos en lenguaje natural [12].

Si bien la estrategia avanza construyendo modelos, su fin no es producir modelos sino alcanzar una profunda comprensión del dominio de la aplicación para definir una solución óptima a través de un sistema de software. Es decir, la estrategia presenta dos

grandes etapas bien distinguibles: una de aprendizaje y la otra de definición. Cuando hay un conocimiento previo del dominio de la aplicación la primera fase se convierte en un proceso confirmatorio.

Los modelos que se utilizan en esta estrategia son:

- 1) un modelo léxico, LEL (Léxico Extendido del Lenguaje) [16], el cual describe el vocabulario utilizado en el dominio de la aplicación,
- 2) un modelo organizacional que describe los procesos actuales del negocio utilizando [17], denominado Escenarios Actuales,
- 3) un modelo organizacional que describe los procesos del negocio proyectados en base al sistema de software a desarrollar [18] [19], denominado Escenarios Futuros.

Los Escenarios Futuros tienen empotrados los requisitos del software. El documento de especificación de requisitos de software, ERS, se puede obtener a partir de estos Escenarios Futuros. El fin de la estrategia utilizada es alcanzar una profunda comprensión del dominio de la aplicación para definir una solución óptima para los clientes/usuarios, a través de un sistema de software.

Desde el punto de vista epistemológico el presente proyecto se puede catalogar como una investigación esencialmente exploratoria, en la que se aspira a detectar: i) los principales factores que contribuyen a definir los mecanismos que permitirían construir rastros entre los diferentes modelos del proceso de requisitos, ii) los requisitos de una herramienta de apoyo al proceso de requisitos y iii) evidencias sobre las debilidades del proceso. Es de esperar que estos resultados puedan luego ser contrastados en investigaciones confirmatorias y que además permitan establecer las bases que permitan diseñar una o más herramientas para facilitar la tarea de construcción y manipulación de los rastros entre modelos.

Si bien se ha procurado y se procurará mantener la máxima generalidad posible, los resultados del proyecto estarán afectados por la estrategia utilizada, consistente en el estudio de un caso. Estas limitaciones están atemperadas por la notoria amplitud del caso y por la continua revisión de los resultados obtenidos, en el marco de la biblioteca de casos estudiados previamente, la que supera largamente la cantidad de 100.

3. RESULTADOS OBTENIDOS/ESPERADOS

Con los trabajos preliminares que sirven de antecedentes al presente proyecto se han podido detectar algunas regularidades que posiblemente se transformen en conclusiones. Estas conclusiones permitirán elaborar hipótesis a ser consideradas en futuros proyectos, los que tendrán carácter de confirmatorios.

Concretamente, hasta el momento se ha descubierto que:

- 1) El vocabulario utilizado en el Universo de Discurso contiene una cantidad importante de símbolos compuestos por frases construidas por sustantivos compuestos, sustantivos adjetivados y frases verbales entre otros. Usualmente la cantidad de estos símbolos supera ampliamente la cantidad de símbolos constituidos por palabras aisladas.
- 2) Las frases verbales suelen estar asociadas semánticamente con objetos. Sin embargo, la forma resultante de estas frases verbales hace que su uso sea prácticamente antagónica con la combinación en una sola sentencia de la frase verbal y el objeto. En otras palabras, una oración suele ser perfectamente no ambigua en su contexto, pero omitir uno de los dos (el verbo o el objeto) o empotrar el objeto en el verbo. Este fenómeno hace que un ingeniero de requisitos, que no ha recibido indicaciones precisas al respecto, tienda a omitir

uno de los dos símbolos. Es importante tener en claro que ninguna de las heurísticas utilizadas hasta el momento mencionan este fenómeno. En un recuento sin validez estadística se ha encontrado que aproximadamente se omiten entre el 30 y el 40% de los símbolos involucrados en esta relación (verbos y objetos). Usualmente se registran más los objetos que los verbos, en una relación groseramente aproximada de 4 a 1. Claramente la existencia de estas omisiones hará que algunos aspectos de las características del Universo de Discurso no se perciban y esto hará que se omitan Escenarios Actuales y posiblemente también Escenarios Futuros. Se desconoce la importancia de estas omisiones las que posiblemente sean atemperadas, en alguna medida, a lo largo de todo el proceso, ya que se consulta el Universo de Discurso en varias actividades del proceso de requisitos, especialmente durante la validación de los diferentes modelos.

- 3) Parece ser que los símbolos compuestos provenientes de la concatenación de sustantivos también ocultan otros símbolos, especialmente aquellos pertenecientes a jerarquías taxonómicas. Sin embargo aún no se ha tenido oportunidad de tener una apreciación más concreta del fenómeno.
- 4) Se ha verificado, también sin comprobación cuantitativa, que la creación automática o semi-automática de rastros se verá facilitada si se procesa todo el conocimiento elicitado de una misma fuente de información ‘en bloque’, significando con esto que sería conveniente agotar el registro del conocimiento capturado de una misma fuente de información, antes de proceder a considerar otra fuente

de información. En esta última afirmación se está haciendo un uso abusivo de ‘fuente de información’ ya que se pretende incluir también la información que se extrae de un modelo para construir otro. En particular en lo que se refiere al LEL:

- a) Ocurre que aunque más de una fuente de información contribuya a completar la información correspondiente a un determinado ítem, esto hace que sea preferible que la gestión de los incrementos se produzca en cada símbolo del LEL y no en las fuentes de información. Esto haría que se incremente la importancia de la actividad de verificación.
- b) Una conclusión aparentemente muy importante surge del hecho que se estima que será necesario modificar la forma de construir el LEL, en virtud que la actividad ‘creación de la lista de símbolos candidatos’ recomendada en la heurística actual es contraproducente, tanto para el registro de rastros como para la reducción de las omisiones.

Todos estos resultados preliminares son sin lugar a duda valiosos porque de confirmarse permitirán elaborar heurísticas más apropiadas, pero su mera observación crítica hace evidente que los mismos constituyen un conjunto semánticamente poco conexo de patrones. Se espera que estos patrones se constituyan en una etapa intermedia en la detección de reglas más abarcativas. Este es el núcleo del presente proyecto.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto de la UNLaM participan cuatro investigadores. En el proyecto de la UNO participan en este tema otros dos investigadores, uno de ellos en formación.

La línea de investigación presentada aquí

es parte directa de las tesis de maestría de la Lic. Renata Guatelli y de las tesis de doctorado de la Mg. Gladys Kaplan y la Ing. Andrea Vera, y colabora con la tesis de doctorado de la Lic. María Pepe.

5. BIBLIOGRAFÍA

- [1] Gotel, O., C., Z., Finkelstein, ACW (1994) An analysis of the requirements traceability problema, ICRE'94. First IEEE International Conference on Requirements Engineering. IEEE Computer Society Press, Colorado Springs, pp. 94-101.
- [2] Palmer J., D., (1997) Traceability. Software Engineering, M. Dorfman y R.H. Thayer Editores. IEEE Computer Society Press. 1996. Pág. 266 a 276. Reimpreso en: Software Requirements Engineering. R.H. Thayer y M. Dorfman Editores. IEEE Computer Society Press. 2º Edición. Los Alamitos, CA. pp. 364-374.
- [3] Pinheiro F., A., C., (2004) Requirements Traceability. Perspectives on Software Requirements. Kluwer Academic Publishers. Capítulo 5, pp. 91-113.
- [4] Cleland-Huang, J., Chang, C., K., Christensen, M., (2013) Event-based traceability for managing evolutionary change. IEEE Trans. Softw. Eng. Vol. 29. N° 9, pp. 796-810.
- [5] De Lucia, A., Fasano, F., Oliveto, R., Tortora, G., (2007) Recovering Traceability Links in Software Artifact Management Systems using Information Retrieval Methods, ACM Transactions on Software Engineering and Methodology. Vol.16. N° 4, pp. 13.1-13.50.
- [6] Pruski, P, Lohar, S., Goss, W., et al (2015) TiQi: answering unstructured natural language trace queries, Requirements Engineering, Vol. 20, Nro. 215, pp. 215-232.
- [7] Mahmoud, A., Niu, N., (2015) On the role of semantics in automated requirements tracing. Requirements Engineering, Vol. 20, Nro. 281, pp. 281-300.
- [8] Leite, J., C., S., P., Doorn, J. H., Kaplan, G., Hadad, G., D., S., Ridao, M., N., (2004) Defining System Context using Scenarios, Perspectives on Software Requirements, Kluwer Academic Publishers, cap.8, pp. 169-199.
- [9] Doorn, J., H., Ridao, M., N., (2003) Completitud en Glosarios: Un estudio experimental, en WER'03, International Workshop on Requirements Engineering, pp. 317-328.
- [10] Ridao, M., N., Doorn, J., H., (2006) Estimación de Completitud en Modelos de Requisitos Basados en Lenguaje Natural, WER'06, Internacional Workshop on Requirements Engineering, pp. 146-157.
- [11] Doorn, J., H., Ridao, M., N., (2008) Completeness Concerns in Requirement Engineering”, en “Encyclopedia of Information Science and Technology, Second Edition”.
- [12] Hadad, D., S., Litvak, C., S., Doorn, J., H., Ridao, M., N., (2015) “Dealing with Completeness in Requirements Engineering”, en “Encyclopedia of Information Science and Technology, Third Edition”.
- [13] Loucopoulos, P., Karakostas, V. (1995) System Requirements Engineering. McGraw-Hill, Londres.
- [14] Rolland, C., Souveyet, C., Ben Achour, C., (1998) Guiding Goal Modeling Using Scenarios, IEEE TSE, Vol. 24, Nro. 12, pp. 1055-1071.
- [15] Neill, CJ., Laplante, PA. (2003) Requirements Engineering: The State of the Practice, IEEE Software, Vol. 20, Nro. 6, pp. 40-45.
- [16] Leite, JCSP, Franco, APM, (1993) A Strategy for Conceptual Model Acquisition, IEEE First International Symposium on Requirements Engineering, RE'93, IEEE Computer Society Press, pp. 243-246, Los Alamitos, EEUU.
- [17] Leite, JCSP, Hadad, GDS, Doorn, JH, Kaplan, GN (2000) A Scenario Construction Process. Requirements Engineering Journal. Vol.5, N° 1, pp. 38-61.
- [18] Hadad GDS (2008) Uso de Escenarios en la Derivación de Software. Tesis Doctoral, Facultad de Ciencias Exactas, Universidad Nacional de La Plata. Argentina. 2008.
- [19] Hadad, GDS, Doorn, JH, Ridao, MN, Kaplan GN (2009) Facilitando la asignación de Prioridades a los Requisitos, WER'09 – 12th Workshop on Requirements Engineering, ISBN: 978-956-319-941-3, Universidad Técnica Federico Santa María, pp. 75-84. Valparaíso. Chile.

Evolución de los Factores Situacionales durante el Proceso de Requisitos

Viviana A. Ledesma¹, Graciela D.S. Hadad^{1,2}, Jorge H. Doorn^{1,2}, Juan P. Mighetti¹,
Nicolás A. Bedetti¹, María Celia Elizalde²

¹DIIT, Universidad Nacional de la Matanza

²Escuela de Informática, Universidad Nacional del Oeste

{vledesma, ghadad}@unlam.edu.ar, jdoorn@exa.unicen.edu.ar, {jmighetti, nbedetti}@unlam.edu.ar,
melizalde@uno.edu.ar

RESUMEN

La literatura presenta una diversidad de procesos de software, que dependen en gran medida del tipo de software a construir y de la envergadura del mismo. La selección del proceso más adecuado y la adaptación a un contexto específico suele ser a su vez fuertemente dependiente de la experiencia de los desarrolladores. Sin embargo, condicionar el análisis de los factores situacionales de esta manera es una restricción innecesaria. Por otra parte, muchos de estos factores impactan en el proceso de software desde sus etapas iniciales. Es por ello que es beneficioso realizar una adaptación lo más temprana posible, en función de factores observables en el contexto. Es decir, considerar los factores situacionales desde el proceso de requisitos mismo. A partir de varios casos reales donde se aplicó un proceso de requisitos, se inició la evaluación de dichos factores, observándose que algunos de ellos son difíciles de establecer con precisión en los primeros pasos del proceso o que por su naturaleza cambian a lo largo del proceso. Conocer qué factores son más proclives a evolucionar por mejora de la estimación inicial o por cambios genuinos permitirá una mejor adaptación del proceso de requisitos.

Palabras clave: Ingeniería de Requisitos, Proceso de Requisitos, Ingeniería de Métodos Situacional, Variabilidad de Procesos.

CONTEXTO

La línea de investigación que se presenta es parte de los proyectos de investigación “Adaptación Dinámica de un Proceso de

Requisitos” del programa CYTMA2 de la Universidad Nacional de La Matanza (UNLaM) y “Tratamiento de los Factores Situacionales y la Completitud en la Ingeniería de Requisitos” de la Universidad Nacional del Oeste (UNO).

1. INTRODUCCIÓN

Implementar una gestión adecuada de los procesos permite a las organizaciones no solo reducir esfuerzos, sino además obtener ventajas competitivas frente a otras organizaciones. Esto no resulta sencillo especialmente cuando se trabaja en ambientes extremadamente dinámicos, los cuales conllevan un alto nivel de variabilidad. Es por ello que los constantes cambios que surgen en dichos entornos crean la necesidad de adaptar continuamente estos procesos. Adaptarse y soportar estos cambios resulta clave para el éxito de las organizaciones [1].

El análisis de variabilidad de procesos involucra establecer qué partes varían (puntos de variación), las causas de esa variación, las alternativas posibles de resolución (flujos variantes) y el mecanismo de resolución de la variabilidad (aplicando operaciones de adición, eliminación, reemplazo o parametrización de partes del proceso) [2, 3].

Es así que la variabilidad puede interpretarse como planear o anticiparse al cambio, en lugar de improvisar cada vez que surge una necesidad o una situación no prevista [3].

Existe un gran número de trabajos de investigación relacionados al análisis de la variabilidad, que ofrecen numerosas propuestas para dominios específicos. Un

campo de estudio es la representación de la variabilidad en los Modelos de Proceso de Negocio [2, 4]. Bajo este enfoque el análisis de la variabilidad consiste en la definición de caminos alternativos de ejecución en un flujo de trabajo, donde es posible indicar qué partes del proceso de negocio permanecen abiertos al cambio o no totalmente definidas, soportando diferentes versiones del mismo proceso, dependiendo del uso proyectado o el contexto de ejecución [5]. En las Líneas de Producto Software también se aplica la gestión de variabilidad [6]. Una técnica que la pone en práctica es la aproximación Base-Variación-Resolución, que propone la construcción de un modelo base, un modelo de variación y un modelo de resolución, para definir de forma explícita las partes del modelo que son fijas, las que pueden cambiar, y las condiciones que harán que éstas cambien [7]. Basándose en la visión de Líneas de Producto Software, se da sustento al concepto de Líneas de Proceso Software [2, 8], facilitando el ajuste sistemático del proceso, incluyendo la reutilización y evolución de forma planificada.

Por lo anterior, se hace evidente que la variabilidad es un factor clave que aplica a muchos aspectos del desarrollo de software en general [9]. En ese sentido, se han propuesto distintas técnicas para modelar la variabilidad que van desde diagramas de actividad de UML [10], hasta otros más específicos tales como la notación para el Modelado de Procesos de Negocio [11], la Especificación de Recursos Reutilizables [12] que define de manera estándar la información asociada a un activo o recurso, incluyendo las guías necesarias para facilitar la manipulación y reutilización del mismo, o el Metamodelo de Ingeniería de Procesos de Software (SPEM) [13] que especifica un lenguaje para modelar familias de procesos de software.

Desde otra visión del modelado de procesos, la Ingeniería de Métodos se enfoca en el diseño de métodos mediante la composición de fragmentos o bloques de proceso existentes, con el fin de mejorar la productividad de los mismos [14]. A partir de esta disciplina, surgió la Ingeniería de

Métodos Situacional (IMS), enfocada en el diseño de métodos de desarrollo de software adaptados a situaciones particulares [15, 16]. La situación se describe a través de la estimación de un conjunto de factores, en base a los cuales se determina la adaptación a realizar del proceso de software [17]. Estos factores muestran tanto peculiaridades del contexto de aplicación como del proyecto de software [18]. Por lo tanto, un proceso se construye para una situación dada a partir de la combinación de bloques de proceso predefinidos y en función de los factores situacionales identificados, donde existen bloques comunes a cualquier situación y otros variantes [16].

La IMS no solo considera el desarrollo de componentes de proceso sino también componentes de producto, e incluso componentes que ensamblan tanto proceso como producto (bloques mixtos) [19]. La ventaja que conlleva el enfoque de uso separado de bloques de proceso y bloques de producto sobre los bloques mixtos, es que el primero brinda una mayor flexibilidad en el proceso de adaptación, ya que varios bloques de proceso podrían asociarse a los mismos bloques de producto [15, 19].

En general, las técnicas que aplican la IMS coinciden en los siguientes pasos para la construcción del nuevo método:

- Caracterizar el proyecto, a partir de definir la situación.
- Seleccionar los fragmentos existentes que conformarán el método.
- Ensamblar los fragmentos de modo tal que se obtenga el método para ese proyecto específico.

En el proceso de IMS propuesto en [16], el ingeniero de métodos selecciona los bloques del proceso base y construye un proceso específico, en función de los factores situacionales y las guías de construcción existentes.

Han surgido algunas propuestas desde la IMS que abordan el diseño y mejora de procesos de Ingeniería de Requisitos (IR), adaptándolos a las características particulares de un proyecto. Algunas propuestas están orientadas a definir un proceso de IR

utilizando frameworks, ya sea que existan en el mercado o propuestos por los autores, partiendo de una base de conocimiento sobre actividades y técnicas de IR, que sirve de apoyo para obtener un proceso acorde a las necesidades del proyecto [20, 21]. Varias propuestas se enfocan exclusivamente en la selección de las técnicas de elicitación que mejor se adapten al contexto situacional [22, 23]. Otros autores presentan modelos para adaptar procesos de IR para proyectos de determinada naturaleza o alcance, tales como desarrollo de software global o métodos ágiles, tomando algunos principios de la IMS para flexibilizar el proceso, incluyendo otros atributos contextuales [24, 25]. En otros trabajos [26, 27] se han aplicado algunos principios de la IMS para definir procesos de IR adaptables, mediante el uso de componentes modulares existentes, incluso combinando nociones de Variabilidad de Procesos y de Líneas de Producto Software.

En estos trabajos, se aplican en mayor o menor medida algunos conceptos de la IMS, proponiendo factores situacionales que son evaluados para definir el contexto, donde algunos enfoques se basan en la reutilización de componentes para conformar el nuevo proceso de IR. Sin embargo, en general no presentan propuestas de adaptación dinámica del proceso. Una aproximación a ello es propuesto en [28], donde se enfatiza que el conocimiento acerca del dominio del problema y del dominio del proyecto se va adquiriendo a medida que avanza el proceso de IR, con lo cual en cada iteración se puede mejorar la actividad de elicitación y de esta forma se mejoraría el producto final. En este sentido, Rolland [19] menciona que las características que describen una situación pueden cambiar a lo largo del proceso, pudiendo ser necesario readaptar el proceso inicialmente construido para esa situación.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A lo largo de dos décadas en varios proyectos de investigación, se ha desarrollado un proceso de IR basado en modelos en lenguaje natural [29], que ha sido difundido

en cursos de grado y posgrado, y puesto en práctica en diversos proyectos de software de mediana a gran envergadura en la industria. Aún cuando este proceso puede considerarse suficientemente maduro, todavía requiere mejoras en algunos aspectos. Una mejora apunta a la flexibilización y optimización del proceso, por lo que se propone mejorar la aplicación de la IMS incorporando nociones de análisis de variabilidad de procesos, y utilizando el proceso existente y las estrategias ya desarrolladas para su adaptación estática, como base para encarar su adaptación dinámica. Esto involucra tomar decisiones referidas a qué artefactos de requisitos construir, qué actividades del proceso realizar y qué técnicas específicas aplicar en cada actividad, de manera de lograr una mayor productividad del proceso, y procurando una mejor calidad de los requisitos del software, considerando la posibilidad que los factores situacionales hayan evolucionado o hayan sido estimados incorrectamente. Entender las decisiones adoptadas en la aplicación de una variante específica del proceso de IR permitirá establecer mejoras en adaptaciones posteriores.

Manejar alternativas en el proceso de IR implica la existencia de bloques de proceso comunes independientes de la adaptación, bloques de proceso alternativos y otros bloques de proceso con alguna variabilidad interna menor. Se asume que la variabilidad en los procesos también podrá influir en los productos, pudiendo entonces tenerse modelos alternativos (por ejemplo, diferentes estructuras para el documento de especificación de requisitos) o familia de modelos (por ejemplo, un conjunto de escenarios vinculado a un glosario del universo de discurso, o vinculado a un glosario que describe términos del sistema de software, o sin vínculos a ningún glosario).

Asimismo, ocurre con frecuencia que las condiciones iniciales de un proyecto e incluso el contexto de aplicación se alteran, pudiéndose necesitar una nueva adaptación del proceso [19]. Es por ello que se propone una adaptación continua del proceso mediante

el monitoreo de los factores que pueden afectarlo en determinados hitos, proveyendo guías precisas para su reajuste.

3. RESULTADOS OBTENIDOS/ESPERADOS

A continuación se enumeran los resultados ya alcanzados: i) se ha identificado un conjunto amplio de 25 factores situacionales con sus interacciones y tipo de evolución, incluyendo factores no contemplados inicialmente relativos al desarrollo de software global y a la calidad esperada de los requisitos; ii) se han determinado los puntos de variación del proceso de requisitos y los factores que impactan en cada punto; iii) se han definido preliminarmente los bloques de proceso necesarios para componer el proceso; y iv) se han establecido las reglas de adaptación del proceso en cada punto de variación, las que permiten seleccionar los bloques que conformarán el proceso para la situación específica. Con base en estas definiciones, se diseñó una versión preliminar de los mecanismos de adaptación dinámica del proceso de requisitos, donde para cada proyecto de software se debe estimar un valor para cada factor situacional, el grado de certeza esperado en su estimación “a priori” y la posibilidad de reevaluarlo en etapas intermedias del proceso de requisitos. Estas definiciones han quedado parcialmente reflejadas en [30].

Se ha iniciado un estudio de la evolución de los factores situacionales durante la aplicación del proceso de IR, disponiéndose de 35 casos de estudio desarrollados entre 2015 y 2017. Se está desarrollando una herramienta de apoyo para la carga y posterior análisis de los cambios sufridos por los factores situacionales al inicio y al final del proceso de requisitos, combinado con el grado de confiabilidad asumido para el factor. En una etapa muy temprana del análisis, se ha notado que ciertos factores, como complejidad del contexto, grado de reingeniería del proceso de negocio, conflicto de intereses de usuarios y envergadura del proyecto, deben ser observados con especial atención durante el proceso de IR pues suelen

cambiar, mientras que factores como novedad del contexto y nivel de rotación del equipo de desarrollo suelen ser percibidos desde el inicio con mayor precisión.

En las próximas etapas de investigación, se profundizará el análisis de la evolución que sufren los factores situacionales a lo largo del proceso de requisitos desde su estimación inicial a su valoración final. De esta manera se intentará identificar el impacto que valoraciones dudosas en ciertos factores pueden tener en el proceso de requisitos adaptado. Se propone establecer alguna heurística que permita alcanzar un valor más preciso desde el inicio de la planeación de dicho proceso.

Por otro lado, se llevará a cabo la puesta a prueba de la adaptación del proceso de IR en aquellas situaciones consideradas más usuales (dada la diversidad de situaciones posibles) con el fin de comprobar el nivel de impacto alcanzado en la productividad del proceso.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto de UNLaM participan tres investigadores y un alumno becario. En el proyecto de UNO participan en este tema otros dos investigadores, uno en formación.

La línea de investigación presentada aquí es parte directa de la tesis de maestría “Estrategia de Requisitos adaptable según factores de situación” que está desarrollando la Ing. Ledesma en UNLaM. El Ing. Mighetti finalizó su tesis de maestría “Mitigación de amenazas a requisitos en el desarrollo global de software usando LEL y Escenarios” en UNLaM, faltando presentarla para su defensa; esta tesis ha permitido identificar factores situacionales no considerados inicialmente.

5. BIBLIOGRAFÍA

- [1] Wheelen, T.L., Hunger, J.D. (2012) Strategic Management and Business Policy: Toward Global Sustainability, 13° ed., Pearson.
- [2] Schnieders, A., Puhlmann, F. (2006) Variability Mechanisms in E-Business Process Families. Intl Conference on Business Information Systems, pp. 583-601.
- [3] Galster, M., Weyns, D., Tofan, D., Michalik, B., Avgeriou, P. (2014) Variability in Software Systems - A Systematic Literature Review. IEEE

- Transactions on Software Engineering, 40, 3, pp.282-306.
- [4] Groefsema, H., Bulanov, P., Aiello, M. (2011) Declarative Enhancement Framework for Business Processes. Intl Conference on Service-Oriented Computing, pp.495-504. New York: Springer.
- [5] Santos, E., Castro, J., Sánchez, J., Pastor, O. (2010) A Goal-Oriented Approach for Variability in BPMN. 13th Workshop on Requirements Engineering, pp. 17-28.
- [6] Pol'la, M., Buccella, A., Cechich, A., Arias, M. (2014) Un modelo de metadatos para la gestión de la variabilidad en líneas de productos de software. 43 JAIIO, Buenos Aires.
- [7] Haugen, Ø., Øgård, O. (2014) BVR – Better Variability Results. Intl Conference on System Analysis and Modeling: Models and Reusability. SAM 2014, Lecture Notes in Computer Science, vol. 8769, pp. 1-15. Springer.
- [8] Simidchieva, B., Clarke, L., Osterweil, L. (2007) Representing Process Variation with a Process Family. Software Process Dynamics and Agility. ICSP 2007. Lecture Notes in Computer Science, vol. 4470, pp. 109-120. Springer, Berlín.
- [9] Chen, L., Babar, M. (2010) Variability Management in Software Product Lines: An Investigation of Contemporary Industrial Challenges. 14th Intl Conference on Software Product Lines: going beyond, pp. 166-180. Springer-Verlag Berlin.
- [10] Razavian, M., Khosravi, R. (2008) Modeling Variability in Business Process Models Using UML. 5th Intl Conference on Information Technology: New Generations. EEUU: IEEE.
- [11] Delgado, A., Calegari, D. (2017) BPMN 2.0 Based Modeling and Customization of Variants in Business Process Families. XLIII CLEI, Córdoba.
- [12] Pacini, K., Braga, R. (2015) An Approach for Reusing Software Process Elements based on Reusable Asset Specification: A Software Product Line Case Study. 10th Intl Conference on Software Engineering Advances, pp. 200-206. Barcelona.
- [13] Martínez-Ruiz, T., García, F., Piattini, M., Munch, J. (2011) Modelling Software Process Variability: An Empirical Study. IET Software, 5,2, pp.172-187.
- [14] Brinkkemper, S. (1996) Method Engineering: Engineering of Information Systems Development Methods and Tools. Information and Software Technology, 38, 4, pp. 275-280.
- [15] Henderson-Sellers, B., Ralyté, J. (2010) Situational Method Engineering: State-of-the-Art Review. Journal of Universal Computer Science, 16, 3, pp. 424-478.
- [16] Henderson-Sellers, B., Ralyté, J., Agerfalk, P., Rossi, M. (2014) Situational Method Engineering, Springer-Verlag Berlin Heidelberg.
- [17] Khan, H.H., bin Mahrin, M., bt Chuprat, S. (2014) Factors for Tailoring Requirement Engineering Process: A Review. Intl Journal of Software Engineering and Technology, 1,1, pp.7-18.
- [18] Bucher, T., Klesse, M., Kurpjuweit, S., Winter, R. (2007) Situational Method Engineering. Situational method engineering: fundamentals and experiences, Springer US, pp. 33-48.
- [19] Rolland, C. (2008) Method engineering: towards methods as services. Making Globally Distributed Software Development a Success Story, Springer Berlin, pp. 10-11.
- [20] Zowghi, D., Firesmith, D., Henderson-Sellers, B. (2005) Using the OPEN Process Framework to Produce a Situation-Specific Requirements Engineering Method. 1st Intl Workshop on Situational Requirements Engineering Processes, pp.59-74, Paris.
- [21] Jiang, L., Eberlein, A. (2008) A Framework for Requirements Engineering Process Development (FRERE). 19th Australian Conference on Software Engineering.
- [22] Hickey, A., Davis, A. (2003) Elicitation technique selection: how do experts do it? 11th IEEE Intl Requirements Engineering Conference, pp.169-178.
- [23] Carrizo, D., Dieste, O., Juristo, N. (2008) Study of elicitation techniques adequacy. 11th Workshop on Requirements Engineering, pp. 104-114.
- [24] Bakhat, K.A., Sarwar, A.A., Motla, Y.H.B., Akhtar, M.C. (2015) A Situational Requirement Engineering Model for an Agile Process. Bahria University Journal of Information & Communication Technology, 8, 1, pp. 21-26.
- [25] Khan, H., bin Mahrin, M., Mali, M. (2016) Situational Requirement Engineering Framework for Global Software Development: Formulation and Design. Bahria University Journal of Information & Communication Technologies, 9, 1, pp. 74-84.
- [26] Coulin, C., Zowghi, D., Sahraoui, A. (2006) A Situational Method Engineering Approach to Requirements Elicitation Workshops in the Software Development Process. Software Process: Improvement and Practice, 11, 5, pp. 451-464.
- [27] Jafarinezhad, O., Ramsin, R. (2012) Development of Situational Requirements Engineering Processes: A Process Factory Approach. 36th IEEE Intl Conference on Computer Software and Applications, pp.279-288.
- [28] Hickey, A.M., Davis, A.M. (2004) A Unified Model of Requirements Elicitation. Journal of Management Information Systems, 20, 4, pp.65-84.
- [29] Leite, J., Doorn, J., Kaplan, G., Hadad, G., Ridao, M. (2004) Defining System Context using Scenarios. Perspectives on Software Requirements, Kluwer Academic Publishers, cap.8, pp.169-199.
- [30] Hadad, G.D.S., Doorn, J.H., Ledesma, V.A. (2018) Dynamic Situational Adaptation of a Requirements Engineering Process. Encyclopedia of Information Science and Technology, 4º ed., IGI Global, EEUU, cap. 646, pp. 7422-7434.

Evaluación de Compatibilidad y Complejidad para el Reuso de Servicios

Andres Flores, Alejandra Cechich, Martin Garriga,
Marcelo Moyano, Alan De Renzis, Diego Anabalon, Franco Corgatelli
Grupo de Investigación en Ingeniería de Software del Comahue (GIISCO)
<http://giisco.uncoma.edu.ar>
Departamento de Ingeniería de Sistemas – Facultad de Informática
Universidad Nacional del Comahue
Buenos Aires 1400, (8300) Neuquén
Contacto: [andres.flores, alejandra.cechich]@fi.uncoma.edu.ar

Resumen

El reuso de artefactos software brinda oportunidades para proveedores y clientes, tanto para acelerar el proceso de desarrollo de software como para establecer oferta de productos reusables. El paradigma de Computación Orientada a Servicios (SOC), promueve el desarrollo de aplicaciones distribuidas en ambientes heterogéneos, que son construidas ensamblando o componiendo servicios reusables, que se publican a través de una red y se acceden mediante protocolos específicos. SOC ha sido ampliamente adoptado bajo su implementación con la tecnología de Servicios Web, que provee flexibilidad de ejecución remota que oculta las plataformas específicas de ejecución y permite descentralizar los procesos de negocios. SOC requiere la publicación de servicios en un registro (UDDI de acuerdo a Servicios Web), los cuales luego son identificados y evaluados para una aplicación en desarrollo. Sin embargo, aún este proceso necesita métodos exhaustivos y eficientes, tanto para identificación como para selección de servicios, en el cual se puede considerar la aplicación de técnicas de Pruebas de Software y el uso de dos conceptos actuales: Orquestación y Coreografía de servicios.

Palabras Clave: Ingeniería de Software basada en Reuso – Software Orientado a Servicios – Servicios Web – Calidad de Software – Verificación y Validación.

Contexto

La línea presentada se inserta en el contexto de los siguientes proyectos y acuerdos de cooperación:

- 04/F009: “Reuso Orientado a Servicios – Parte II”, sub-proyecto del Programa “Desarrollo de Software basado en Reuso – Parte II”. Financiado por UNCo. (2017-2020).
- PIP 2017-2019 11220170100951CO: “Construcción de Líneas de Productos Software guiada por Estándares de Dominio”. Financiado por CONICET.
- Acuerdo de Cooperación con ISISTAN-CONICET, UNICEN, Tandil.
- Acuerdo de Cooperación con el Grupo Alarcos, Escuela Superior de Informática, Universidad de Castilla-La Mancha, España.

Introducción

Actualmente la industria de software observa cómo mediante el reuso de software se puede alcanzar un proceso de desarrollo de software acelerado y confiable al basarse en artefactos software

que ya han sido probados en diferentes contextos de aplicación. Para ello se adopta el concepto denominado “tercerización”, por medio del cual se acuerdan contratos para adquisición y provisión de artefactos software reusables y se establecen relaciones comerciales entre vendedores y clientes. Por lo tanto, desde el punto de vista de un cliente implica la posibilidad de acelerar el desarrollo de un producto software para reducir el lanzamiento al mercado, y desde el punto de vista de un proveedor implica la posibilidad de observar sus productos con una perspectiva nueva que los coloque dentro del mercado de artefactos reusables.

Un paradigma que promueve altamente el reuso de software se denomina Computación Orientada a Servicios (SOC), donde la funcionalidad a ser reusada adopta la forma de servicios, o unidades lógicas que presentan entornos heterogéneos de ejecución y pueden ser ensambladas para formar otras unidades lógicas de mayor nivel de abstracción que resuelvan (directamente o en parte) los procesos de negocios para un contexto de aplicación [SH05,PTDL07]. El paradigma SOC encontró una plataforma potencial de aprovechamiento mediante la Web, desde donde se desarrolló la tecnología de Servicios Web [NSS03, Wetal05], cuya base es el esquema estándar XML y el Lenguaje de Descripción de Servicios Web (WSDL), que facilitan ampliamente el desarrollo y mantenimiento de especificaciones formales de servicios. Así el paradigma SOC bajo la implementación con Servicios Web ha logrado su amplia adopción en la industria, principalmente bajo la flexibilidad de ejecución remota que permite a las compañías descentralizar aún más sus procesos de negocios y la ventaja de que las plataformas específicas de ejecución se encuentran ocultas, por lo

cual no se requiere de inversiones adicionales en tecnología (incluyendo costos y esfuerzo de aprendizaje) al adquirir funcionalidad de terceras partes. El beneficio que la tecnología de Servicios Web provee al paradigma SOC se ha denominado “relación sin responsabilidad”, donde una aplicación cliente no requiere asumir cómo se ha implementado el servicio con el que se comunica. Sin embargo, los proveedores de servicios tienen la responsabilidad de evaluar la calidad de los productos ofrecidos como servicios y los consumidores de servicios a su vez deben ser capaces de identificar tal calidad que influirá sobre las aplicaciones en desarrollo. Para ello, las estrategias de Pruebas de Software deben ajustarse a este contexto específico donde los servicios están acordados como cajas negras que sólo permite evaluar el comportamiento y cualidades observables externamente [BDN10, Z08].

El funcionamiento concreto del paradigma SOC se basa en la Arquitectura orientada a Servicios (SOA) [SH05] que se encuentra compuesta por tres actores principales: un proveedor, un consumidor y un registro de servicios; donde el proveedor desarrolla y publica servicios en el registro, para que luego el consumidor busque servicios y establezca una comunicación con el proveedor. Sin embargo, la búsqueda de servicios publicados en un registro UDDI (según la tecnología de servicios Web) [OASIS04], en general requiere invertir un esfuerzo considerable para distinguir servicios candidatos que satisfagan los requerimientos de la aplicación cliente [NSS03, Wetal05]. En particular, cuando varios candidatos ofrecen funcionalidades similares se requieren métodos eficientes de selección de servicios que discriminen tanto aspectos funcionales como no-funcionales, considerando además las

interacciones válidas para un servicio candidato en función de los procesos de negocio que implementará la aplicación cliente. En particular el ensamblaje de servicios considerando procesos de negocio e interoperabilidad de servicios plantea el uso de dos conceptos de reciente investigación: Orquestación y Coreografía de servicios [P03, Wetal05]. El primero relacionado a una aplicación particular que describe un proceso de negocios específico, y el segundo relacionado a las interacciones válidas que pueden ocurrir entre distintos servicios predestinados a intervenir en una colaboración. En este contexto se cuenta actualmente con diversos lenguajes de descripción y frameworks de ejecución, tales como BPEL4WS [OASIS07], y WSCDL (Web Services Choreography Description Language) [W3C05].

Líneas de Investigación, Desarrollo e Innovación

El perfil de esta línea puede definirse en base a las actividades de investigación y transferencia, a las que da soporte el grupo GIISCo. Los temas específicos consideran los desafíos diferentes asociados al crecimiento de la Tecnología de la Información y las Comunicaciones. Actualmente, abordamos los siguientes aspectos:

- Complejidad y legibilidad de servicios
- Compatibilidad y selección de servicios.
- Adaptación y Composición de servicios.
- Testing de servicios.
- Herramientas para evaluación, selección, composición y testing de servicios.
- Definición de aplicaciones en dominios específicos.

Resultados y Objetivos

En [FCGMRAC17] hemos enumerado una serie de contribuciones anteriores. Durante el año 2017, hemos profundizado la investigación en aspectos de *evaluación de compatibilidad y complejidad de servicios*, generando métodos y herramientas enfocados en las interfaces y comportamiento dinámico de los servicios. Este avance se ha efectuado en colaboración con investigadores de ISISTAN (UNICEN) [AFMZM17, GRFCZ17, MZMAF17a, MZMAF17b, RGFCMZ17a, RGFCMZ17b].

Las líneas de investigación convergen en el tratamiento del desarrollo de software basado en el reuso de servicios desde la perspectiva de las aplicaciones orientadas a servicios. Una aplicación orientada a servicios implica una solución de negocio que consume servicios de uno o más proveedores y los integra en un proceso de negocio [SW04]. Además puede verse como una aplicación basada en componentes que integra dos tipos de componentes: internos localmente empotrados en la aplicación, y externos estática o dinámicamente enlazados a algún servicio [CMZC14]. No solamente se enfocará en el reuso de servicios individuales, sino también en la composición de servicios como forma de tercerizar una funcionalidad. Se adoptará la visión de proceso de negocio para la definición de comportamiento, donde se aplicará testing de servicios para una evaluación dinámica. Se complementará el modelo de selección y composición de servicios mediante las últimas plataformas y avances tecnológicos incluyendo semántica y estandarización. Se prevee la aplicación de estos modelos y las herramientas de soporte a dominios específicos, con particular énfasis en aquellos que requieran rigurosidad como aporte de validación efectiva. La visión

de esta línea de investigación se resume en:

“Definir técnicas y herramientas para la mejora del desarrollo de software, en función del reuso de servicios web. La definición de modelos de identificación y selección de servicios, y la posibilidad de composición de servicios”.

Formación de Recursos Humanos

Este proyecto se compone de 11 investigadores, entre los que se cuentan docentes y estudiantes del Grupo GIISCo de UNComa y asesores externos. Algunos de los docentes–investigadores se encuentran realizando carreras de postgrado. Se cuenta actualmente con 3 doctores (1 investigador adjunto CONICET, y 1 investigador asistente CONICET), 2 doctorandos (becarios CONICET), 1 maestrando, y 1 becario EVC-CIN entre los miembros del proyecto. Dirección de Tesis de Grado durante 2017: 3 tesis.

Referencias

- [AFMZM17] Anabalón, D., Flores, A., Mateos, C., Zunino, A., Misra, S. (2017). *Controlling Complexity of Web Services Interfaces through a Metrics-driven Approach*. ICCNI'17, International Conference on Computing Networking and Informatics. pp. 1-9. IEEE Computing Society Press. Ota, Nigeria.
- [BDN10] Baresi, L.; Di Nitto, E. (2010). *Test and Analysis of Web Services*. Springer.
- [CMZC14] Crasso, M., Mateos, C., Zunino, A., Campo, M. (2014). *EasySOC: Making Web Service Outsourcing Easier*. International Journal on Information Sciences, 259: 452–473.
- [FCGMRAC17] Flores, A., Cechich, A., Garriga, M., Moyano, M., De Renzis, A., Anabalón, D., Corgatelli, F. (2017). *Reuso Orientado a Servicios: Compatibilidad y Complejidad de Servicios*. WICC'17, Workshop de Investigadores en Ciencias de la Computación. pp. 622-626. Buenos Aires.
- [GRFCZ17] Garriga, M., De Renzis, A., Flores, A., Cechich, A., Zunino, A. (2017). *Assessing Web Services Interfaces with Lightweight Semantic Basis*. Computing & Informatics, 36(5):1173-1206. (**Indexed SCI IF JCR2015: 0.504**).
- [MZMAF17a] Mateos, C., Zunino, A., Misra, S., Anabalón, D., Flores, A. (2017). *Managing Web Service Interface Complexity via an OO Metric-based Early Approach*. CLEI Electronic Journal, 20(3): 1-22. (**Latindex**).
- [MZMAF17b] Mateos, C., Zunino, A., Misra, S., Anabalón, D., Flores, A. (2017). *Migration from COBOL to SOA: Measuring the Impact on Web Services Interfaces Complexity*. ICIST'17, International Conference on Information and Software Technologies, pp. 266-279. Springer-Verlag CCIS. Lithuania.
- [NSS03] Nagappan, R.; Skoczylas, R.; Sriganesh, R. (2003). *Developing Java™ Web Services: Architecting and Developing Secure Web Services Using Java*. Wiley Publishing Inc.
- [OASIS04] OASIS Consortium (2004). *UDDI Version 3.0.2*. UDDI Spec Technical Committee Draft, October.
- [OASIS07] OASIS Standard (2007). *Web Services Business Process Execution Language Version 2.0*. <http://docs.oasis-open.org/wsbpel/2.0/OS/wsbpel-v2.0-OS.html>

- [P03] Peltz, C. (2003). *Web Services Orchestration and Choreography*. IEEE Computer, 36(10): 46–52.
- [PTDL07] Papazoglou, M.; Traverso, P.; Dustdar, S.; Leymann, F. (2007). *Service-Oriented Computing: State of the Art and Research Challenges*. IEEE Computer, 40(11): 38–45.
- [RGFCMZ17a] De Renzis, A., Garriga, M., Flores, A., Cechich, A., Mateos, C., Zunino, A. (2017). *Assessing Readability of Web Service Interfaces*. CLEI Electronic Journal, 20(2): 1-23. (**Latindex**).
- [RGFCMZ17b] De Renzis, A., Garriga, M., Flores, A., Cechich, A., Mateos, C., Zunino, A. (2017). *A Domain Independent Readability Metric for Web Service Descriptions*. Computer Standards & Interfaces, 50: 124–141. Elsevier (**Indexed SCI, IF JCR2016: 1,268**).
- [SH05] Singh M.; Huhns, M. (2005). *Service-oriented computing: Key concepts and principles*. IEEE Internet Computing, 9(1): 75–81.
- [SW04] Sprott, D.; Wilkes, L. (2004). *Understanding Service-Oriented Architecture*. The Architecture Journal. MSDN Library. Microsoft Corporation, <http://msdn.microsoft.com/en-us/library/aa480021.aspx>
- [W3C05] W3C Candidate Recommendation (2005) *Web Services Choreography Description Language Version 1.0*. <http://www.w3.org/TR/ws-cdl-10/>
- [Wetal05] Weerawarana, S.; Curbera, F.; Leymann, F.; Storey, T.; Ferguson, D. (2005). *Web Services Platform Architecture: SOAP, WSDL, WS-Policy, WS-Addressing, WS-BPEL, WS-Reliable Messaging, and More*. Prentice Hall PTR.
- [Z08] Zhou, X. (2008). *Testing and Verifying Web Services. From the Researcher's Perspective*. VDM Verlag.

Tendencias en el desarrollo de Aplicaciones para Dispositivos Móviles

Pablo Thomas, Lisandro Delia, Leonardo Corbalan, Germán Cáseres,
Juan Fernandez Sosa, Fernando Tesone, Alfonso Cuitiño, Patricia Pesado

Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – Universidad Nacional de La Plata

50 y 120 La Plata Buenos Aires

Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

{pthomas, ldelia, corbalan, gcaseres, jfernandez, ftesone, acuitino, ppesado}@lidi.info.unlp.edu.ar

Resumen

Se presenta una línea de investigación y desarrollo centrada en el estudio de aspectos de Ingeniería de Software aplicados a las diferentes tendencias en el desarrollo de Aplicaciones para Dispositivos Móviles.

Palabras claves: Plataformas para Dispositivos Móviles - Aplicaciones Nativas- Aplicaciones Híbridas – Aplicaciones Interpretadas – Aplicaciones por Compilación Cruzada – Aplicaciones Web Progresivas – Offline First. – Instant App

Contexto

Esta línea de Investigación forma parte del Proyecto (2018-2021) “*Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso*”, en particular del subproyecto *Ingeniería de Software para escenarios híbridos* del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación de la Nación.

Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa en proyectos financiados por el Ministerio de Ciencia y Tecnología de España y la AECID.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Por otra parte, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado.

Introducción

La computación móvil puede definirse como un entorno de cómputo con movilidad física. Un usuario debe ser capaz de acceder a datos, información u otros objetos lógicos desde cualquier dispositivo en cualquier red, al mismo tiempo que va moviéndose, cambiando su locación geográfica.

El desarrollo de aplicaciones para dispositivos móviles plantea nuevos desafíos originados en las características únicas de esta actividad. La necesidad de tratar con diversos estándares, protocolos y tecnologías de red; las capacidades limitadas, aunque en constante crecimiento, de los dispositivos; las restricciones de tiempo impuestas por un mercado altamente dinámico y la existencia de distintas plataformas de hardware y software son sólo algunas de las dificultades a las que se enfrentan los desarrolladores en esta área.

Las aplicaciones se generan en un entorno dinámico e incierto. En su mayoría se trata de aplicaciones pequeñas, no críticas, destinadas a un gran número de usuarios finales que son liberadas en versiones rápidas para poder satisfacer las demandas del mercado. En otros casos las aplicaciones son de mayor tamaño, algunas incluso pueden funcionar fuera de línea requiriendo sincronización con bases de datos u otro tipo de aplicaciones. En este

último caso se denominan aplicaciones *offline first* y poseen un mayor grado de complejidad, tanto en su desarrollo como en su mantenimiento.

Todas las particularidades previamente mencionadas hacen que el desarrollo de software para dispositivos móviles difiera considerablemente del tradicional. Ello conduce a nuevas prácticas y metodologías que promueven el crecimiento de la Ingeniería de Software como disciplina, acompañando este proceso de desarrollo tecnológico.

La existencia de una enorme competencia en el mercado de las aplicaciones móviles, promovida en parte por la proliferación de dispositivos móviles y el número de negocios que están migrando sus servicios a Internet, revela la importancia que implica para las empresas el posicionamiento en este mercado. Para conseguirlo, es necesario reducir al máximo el tiempo de desarrollo de las aplicaciones y al mismo tiempo hacer que éstas se ejecuten en el mayor número de dispositivos posible.

Este propósito se ve obstaculizado por la excesiva fragmentación de hardware y software existente, originada por el alto número de dispositivos distintos, con sus propios sistemas operativos y plataformas de desarrollo.

La alternativa más costosa para hacer frente a esta problemática consiste en el desarrollo nativo de las aplicaciones en cada una de las plataformas existentes, utilizando el entorno de desarrollo integrado (IDE por sus siglas en inglés), el lenguaje de programación y las herramientas propias de cada plataforma.

Las aplicaciones nativas así desarrolladas poseen un conjunto de características ventajosas entre las que sobresalen el acceso a todas las capacidades del dispositivo (cámara, GPS, acelerómetro y agenda, entre otras), el alto rendimiento, la posibilidad de trabajar sin acceso a Internet y de correr en segundo plano notificando al usuario sólo en caso de requerir su atención. Estas aplicaciones pueden distribuirse a través de las tiendas en línea correspondientes. Sin embargo, el precio de todas estas ventajas es alto: no es posible

reusar el código fuente entre plataformas diferentes, el esfuerzo se multiplica y se elevan los costos de desarrollo, actualización y distribución de nuevas versiones.

El desarrollo multiplataforma se contrapone al nativo y se centra en la reutilización de código. Se procura entonces optimizar la relación costo/beneficio compartiendo la misma codificación entre las versiones para las distintas plataformas. La construcción de aplicaciones Web Móviles constituye un ejemplo representativo de este enfoque. Estas aplicaciones se diseñan para correr dentro de un navegador, se desarrollan con tecnología web bien conocidas (HTML, CSS y JavaScript), no necesitan adecuarse a ningún entorno operativo; su puesta en marcha es rápida y sencilla.

Las desventajas de las aplicaciones Web Móviles recaen sobre su rendimiento. Los tiempos de respuesta se dilatan afectados por la interacción cliente-servidor y las restricciones de seguridad impuestas a la ejecución de código por medio del navegador limitan el acceso a todas las capacidades del dispositivo. Además, al no poseer el *look and feel* de las aplicaciones nativas, resultan menos atractivas para el usuario final.

Las aplicaciones híbridas constituyen otro tipo de desarrollo multiplataforma basado en tecnologías web (HTML, Javascript y CSS) pero que, a diferencia de las anteriores, no son ejecutadas por un navegador. En su lugar, corren en un contenedor web especial con mayor acceso a las capacidades del dispositivo a través de una API específica.

Las aplicaciones híbridas permiten la reutilización de código en las distintas plataformas, el acceso al hardware del dispositivo, y la distribución a través de las tiendas de aplicaciones. Sin embargo conservan algunas de las desventajas de las aplicaciones Web Móviles: la utilización de componentes no nativos en la interfaz perjudica la experiencia de usuario, y la ejecución se ve ralentizada por la carga asociada al contenedor web.

Otro tipo de aplicación multiplataforma lo constituyen las aplicaciones interpretadas, las

cuales son traducidas en su mayor parte a código nativo, mientras que un resto se interpreta en ejecución. Se implementan de forma independiente de las plataformas utilizando diversas tecnologías y lenguajes, tales como Java, Ruby y XML, entre otros.

La obtención de interfaces nativas constituye una de las principales ventajas de este tipo de aplicaciones, y la dependencia total con el entorno de desarrollo el obstáculo más notable. Appcelerator Titanium es el entorno de desarrollo más popular.

Finalmente, las aplicaciones generadas por compilación cruzada también constituyen un tipo de desarrollo multiplataforma. Estas aplicaciones se compilan de manera nativa creando una versión específica de alto rendimiento para cada plataforma destino. Ejemplos de entornos de desarrollo para generar aplicaciones por compilación cruzada son Applause, Embarcadero Delphi XE6 y Xamarin.

Un nuevo concepto ha surgido en los últimos años denominado Aplicaciones Web Progresivas (PWA por sus siglas en inglés). Una PWA es una aplicación web que utiliza las últimas tecnologías disponibles en los navegadores para ofrecer en dispositivos móviles una experiencia lo más parecida posible a la de una aplicación nativa.

Los objetivos que persiguen las PWA son: lograr el mayor rendimiento posible en dispositivos móviles, que la aplicación cargue de manera casi instantánea, que la interfaz de usuario se parezca lo máximo posible a una nativa, que se pueda trabajar sin conexión (*offline first*) y que se puedan enviar notificaciones a los usuarios, como en una aplicación nativa.

Desde 2017 los desarrolladores de Android tienen una nueva opción para hacer llegar sus apps a los usuarios finales. El concepto de *Instant App* permite ejecutar una funcionalidad específica de una aplicación sin necesidad de instalar la App completa.

Líneas de Investigación y Desarrollo

- Enfoques de desarrollo de Aplicaciones para Dispositivos Móviles

- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para dispositivos móviles.
- Aplicaciones Nativas.
- Aplicaciones Web Móviles.
- Aplicaciones Móviles Híbridas (PhoneGap, Sencha Touch, Ionic).
- Aplicaciones Móviles Interpretadas (Appcelerator Titanium, Native Script).
- Aplicaciones Móviles generadas por compilación cruzada (Xamarin, Applause, Embarcadero Delphi XE6).
- *Instant App*, *offline first* y PWA
- Análisis y estudio comparativo de rendimiento, consumo de energía, tamaño de software, entre otros, en los distintos enfoques de Aplicaciones Móviles.
- Experiencia de usuario en Aplicaciones Móviles generadas con distintos enfoques de desarrollo.

Resultados obtenidos/esperados

Los resultados obtenidos/esperados se pueden resumir en:

- Se ha estudiado la manera en que el enfoque de desarrollo utilizado afecta el rendimiento de aplicaciones que realizan procesamiento intensivo. Se extrajeron conclusiones considerando las plataformas iOS y Android junto a varios frameworks de desarrollo multiplataforma.
- Se ha estudiado el impacto que distintos frameworks de desarrollo multiplataforma tienen sobre el tamaño de la aplicación construida. Los resultados obtenidos en este estudio son relevantes dado que el espacio disponible se ha convertido en un recurso crítico para muchos usuarios.
- Se ha desarrollado "*Informática UNLP*", una aplicación móvil para la comunidad de la Facultad de Informática de la Universidad Nacional de La Plata. Es una herramienta de software multiplataforma en continuo crecimiento, que mejora la

comunicación entre alumnos y docentes, facilitando además el acceso a la información relevante sobre la Facultad y sus carreras. Además mediante realidad aumentada permite consultar información de la ocupación de las aulas en tiempo real. Se espera ampliar su funcionalidad desarrollando nuevas características en función del resultado de análisis y revisiones constantes.

- Se ha iniciado un detallado estudio comparativo de distintos enfoques de desarrollo que contempla una extensa lista de características que se presentan de forma particular en cada uno de ellos. Se esperan obtener conclusiones de utilidad para asistir a los desarrolladores en la elección de la opción más ajustada a sus necesidades.
- Se ha iniciado un minucioso estudio sobre la eficiencia energética y su relación con los enfoques de desarrollo. En particular se está analizando la autonomía de las baterías de los dispositivos móviles en aplicaciones con funciones multimedia (acceso a imágenes y reproducción de video) y con alta carga de procesamiento, generadas con diversos enfoques de desarrollo. Los resultados provisorios son alentadores, se esperan alcanzar conclusiones de gran utilidad.
- Entre los objetivos de esta línea de investigación se propone estudiar el alcance de las PWA, analizando limitaciones y ventajas en relación con las Aplicaciones Web Móviles y las Aplicaciones Nativas, haciendo foco en aspectos técnicos y sociales, de comportamiento del usuario y relativos al rendimiento.
- Se plantea analizar las ventajas y desventajas de las "Instant App" de Android. Estudiar y comparar con las aplicaciones nativas tradicionales.
- Esta línea de investigación establece también como objetivo el estudio del concepto "offline first" analizando estrategias de soluciones a cuestiones inherentes a esta tecnología.

- Además, el desarrollo de esta línea de investigación ha de promover el avance sostenido y continuo de la formación de los miembros involucrados en ella.

Formación de Recursos Humanos

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con el proyecto. Además participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

Referencias

1. Hayes, I. S. *Just Enough Wireless Computing*. Prentice Hall Professional Technical Reference. 2002. ISBN:0130994618
2. Spyros Xanthopoulos, Stelios Xinogalos, *A Comparative Analysis of Cross-platform Development Approaches for Mobile Applications*, BCI' 2013, Greece
3. Tracy, K.W., *Mobile Application Development Experiences on Apple's iOS and Android OS*, Potentials, IEEE, 2012.
4. Delia L., Galdamez N., Thomas P, Pesado P., *Un Análisis Experimental de Tipo de Aplicaciones para Dispositivos Móviles*, CACIC 2013, CAECE Mar del Plata, Octubre 2013
5. <http://www.appcelerator.com/>
6. <https://github.com/applause/applause>
7. <https://www.embarcadero.com/es/products/delphi>
8. <http://xamarin.com/>
9. Abrahamsson, P. *Mobile software development -the business opportunity of today*. Proceedings of the International Conference on Software Development, (pp. 20-23). 2005. Reykjavik.
10. <http://devgirl.org/2012/12/04/easy-phonegap-push-notifications-with-pushwoosh/>
11. <http://phonegap.com/>
12. <http://jquerymobile.com/>
13. <http://backbonejs.org/>
14. Digital Possibilities. Mobile Development Frameworks Overview <http://digital-possibilities.com/mobile-development-frameworks-overview/>

15. Markus Falk. Mobile Frameworks Comparison Chart, <http://www.markus-falk.com/mobile-frameworks-comparison-chart/>
16. Anup Kumar y Bin Xie, *Handbook of Mobile Systems Applications and Services*. Editorial CRS Press, ISBN 978-1-4398-0152-9, Año 2012.
17. Sambasivan, D.; John, N.; Udayakumar, S.; Gupta, R., *Generic framework for mobile application development*, Internet (AH-ICI), 2011 Second Asian Himalayas International Conference on Computing & Processing (Hardware/Software).
18. Choi, Y.; Yang, J.-S.; Jeong, J., *Application framework for multi platform mobile application software development*, Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on Computing & Processing (Hardware/Software)
19. Anthony Wasserman , Carnegie Mellon Silicon Valley, *Software Engineering Issues for Mobile Application Development*, 2° Workshop on Software Engineering for Mobile Application Development, MobiCASE '11, Santa Monica, California, USA, October 2011.
20. Jemel, M., Serhrouchni, A. *Content protection and secure synchronization of HTML5 local storage data*. Consumer Communications and Networking Conference (CCNC), 2014 IEEE 11th, Las Vegas, NV, USA.
21. Delia L., Galdamez N., Corbalan L., Thomas P, Pesado P., *Multi-Platform Mobile Application Development Analysis*, IEEE Ninth International Conference on Research Challenges in Information Science, May 2015, Athens, Greece, ISBN 978-1-4673-6630-4
22. Delia L., Galdamez N., Corbalan L., Thomas P, Pesado P., *Un Análisis comparativo de rendimiento en Aplicaciones Móviles Multiplataforma*, CACIC 2015, UNNOBA Junín, Octubre 2015.
23. Delia L., Galdamez N., Corbalan L., Thomas P., Pesado P. *Approaches to Mobile Application Development: Comparative Performance Analysis* SAI Computing Conference (SAI), 2017. Londres, del 18 al 20 de Julio de 2017.
24. Fernández Sosa J., Cuitiño A., Thomas P., Delia L., Caseres G., Corbalán L., Pesado P. *“Informática UNLP” la App de la Facultad de Informática*. CACIC 2017. La Plata, del 9 al 13 de Octubre de 2017.
25. Tim A. Majchrzak, Andreas Bjørn-Hansen, Tor-Morten Grønli, *Progressive Web Apps: the Definite Approach to Cross-Platform Development?*, Proceedings of the 51st Hawaii International Conference on System Sciences j, 2018, ISBN: 978-0-9981331-1-9

Gobernanza Digital. Mejora de procesos de gestión y Calidad de software

Esponda Silvia, Pasini Ariel, Boracchia Marcos,
Calabrese Julieta, Muñoz Rocio, Preisegger Santiago, Pesado Patricia

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 La Plata Buenos Aires
Centro Asociado CIC

(sesponda, apasini, marcosb.jcalabrese, rmunoz, jspreisegger, ppesado) @lidi.info.unlp.edu.ar

Resumen

El III-LIDI (Instituto de Investigación en Informática LIDI) posee un grupo dedicado a la investigación y desarrollo relacionado con la mejora de los procesos de gestión y el aseguramiento de la calidad en procesos de gestión y productos de software.

En este contexto se ha avanzado en conceptos de Gobernanza Digital (estudiando la generación de políticas de prestación de servicios y herramientas de uso en los mismos, que beneficien a la comunidad; en la Mejora de procesos en el ámbito del desarrollo del software, mediante el estudio de diferentes metodologías ágiles y la combinación de herramientas de gestión que asistan al desarrollo de software definiendo, en el ámbito administrativo de la Universidad, guías de trabajo, instructivos y procesos que faciliten la prestación de los servicios; y en el estudio de estándares internacionales de Calidad de Producto y desarrollo de herramientas para asistir a su proceso de evaluación.

Palabras Claves

Ingeniería de Software – Calidad del proceso – Calidad del producto - Normas de Calidad – ISO

Contexto

Esta línea de investigación se enmarca en el proyecto “*Metodologías, técnicas y*

herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso” y en el subproyecto “*Gobernanza Digital. Mejora de Procesos.*” del (2018-2021), en proceso de acreditación como proyecto de incentivo de la UNLP. Asimismo el III-LIDI participa en el proyecto de Innovación y Transferencia en Áreas Prioritarias de la Provincia de Buenos Aires (PIT-AP-BA) “*Herramientas para el desarrollo y la entrega de servicios públicos digitales de acción social para municipios bonaerenses*” en colaboración con el Laboratorio de I&D en Ingeniería de Software y Sistemas de Información de la Universidad Nacional del Sur (UNS); y de proyectos de la Facultad de Informática y de otros organismos.

El Instituto posee diversos acuerdos de cooperación con varias Universidades de Argentina y del exterior y con empresas privadas del sector, interesadas en mejorar sus procesos de desarrollo aplicando mejoras.

Cabe destacar que el Instituto es, desde el año 2015 Centro Asociado de la CIC.

Además, participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Introducción

El uso de diferentes Tecnologías de Información y Comunicación (TICs) ha modificado el desarrollo de un gran

número de actividades cotidianas, principalmente en las relaciones sociales, culturales y económicas.

El concepto de “Nube” como repositorio de información está cada vez más incorporado en la vida diaria. El volumen de información almacenada y la velocidad de procesamiento que se dispone actualmente, permite consumir la información en tiempo real (con técnicas de BigData) para asistir a las personas en sus actividades diarias. Por ejemplo, el asistente de tráfico de Google, que en función de la información que va recolectando de los usuarios en tiempo real, es capaz de ofrecer la ruta óptima para ese momento. Este ejemplo presenta un caso particular vinculado al concepto de ciudades digitales.

Una ciudad digital es aquella en la que, utilizando los recursos propios de la infraestructura de telecomunicaciones y de la informática ya existentes, entre ellas Internet, brinda a sus habitantes un conjunto de servicios digitales a fin de mejorar el nivel de desarrollo humano, económico y cultural de esa comunidad, tanto a nivel individual como colectivo. La infraestructura para brindar estos servicios, tales como redes telefónicas e Internet, se sustenta en los recursos propios de los usuarios e instituciones. Ligado al concepto de ciudad digital, se encuentran los términos gobierno y gobernanza digital.

La gobernanza se lleva a cabo a través de la prestación de servicios públicos. Se define como servicio público a la “actividad llevada a cabo por la Administración o, bajo un cierto control y regulación de ésta, por una organización, especializada o no, y destinada a satisfacer necesidades de la colectividad.” Si dichos servicios son brindados a través del uso de TICs, son considerados servicios públicos electrónicos.[1]

Los servicios públicos electrónicos son realizados por procesos bien definidos e implementados utilizando sistemas de

software. Se hallan diferentes estrategias de enfocar la calidad de estos servicios.

Existe un conjunto de normas relacionadas a la calidad del software que se pueden clasificar en tres grupos explicados a continuación.

El primer grupo está comprendido por las asociadas a la calidad del producto de software, el segundo son normas relacionadas al proceso de desarrollo del software y el tercer grupo, más genérico que los anteriores, son las normas relacionadas con la gestión de la organización desarrolladora de software.

Dentro de la línea de la investigación del proyecto, se destacan los siguientes ejes principales:

1- Mejora de los servicios de gobierno digital en organismos públicos de gobierno

El III-LIDI participa en conjunto con el LISSI-UNS en el proyecto PIT-AP-BA “*Herramientas para el desarrollo y la entrega de servicios públicos digitales de acción social para municipios bonaerenses*” que propone mejorar la eficacia de la implementación de las políticas sociales municipales mediante soluciones informáticas que permitan mejorar la entrega de servicios públicos de acción social.

El proyecto se enfoca en la investigación sobre las características de gobierno abierto y su aplicación en las distintas agencias gubernamentales, sobre herramientas que asistan a las agencias de gobierno en la publicación de sus datos, y sobre la gobernanza móvil, con el objetivo de diseñar soluciones y estrategias para integrar diferentes actividades y utilizar las redes sociales como medio de difusión y de participación, permitiendo mejorar la eficiencia y eficacia de las interacciones gobierno-ciudadanos.

En particular, dentro de la gobernanza digital, se encuentra el uso de diversas tecnologías, para la elección de autoridades mediante el voto (presencial, semipresencial y remoto) de su comunidad. Desde el año 2003 el III-LIDI trabaja en aplicaciones en esta área, entre las cuales se destacan la definición e implementación de estos tres modelos de distintos tipos de votaciones (urnas electrónicas, ambientes de votación, comunicaciones, entre otras),

2- Mejora en los procesos de gestión de la Facultad de Informática

El III-LIDI, en coordinación con el área de Certificación de Calidad de la Facultad de Informática, trabaja en este proyecto, con el objeto de analizar, definir y establecer un plan a ser aplicado a distintos procesos de la Gestión Universitaria.

Desde el año 2011, la Facultad de Informática ha iniciado el camino hacia la certificación de distintos procesos.

El curso de Nivelación a Distancia para el Pre-Ingreso, logró en el año 2012 la certificación IRAM-ISO 9001:2008 del Sistema de Gestión de Calidad (SGC) del "Diseño y realización del curso de Nivelación a Distancia para el Pre-Ingreso a la Facultad de Informática" y en 2015 la Re-certificación

En el 2016 se ha obtenido la certificación IRAM-ISO 9001:2008 del SGC del área de Concursos Docentes: "Llamado a Concursos Docentes Ordinarios", "Concurso Auxiliar Docente Ordinario" y "Concurso Profesor Ordinario" de acuerdo a la ordenanza 179 de la universidad nacional de la plata, y las ordenanzas 303 y 308 de la facultad de informática".

En el 2017 se realizaron las auditorías de seguimiento de los procesos certificados. Durante el año 2018 se realizarán las auditorías de Re-certificación de los procesos bajo la nueva versión de la

norma IRAM-ISO 9001:2015, con la modificación en el alcance del Pre-Ingreso.[2]

3- Mejora de los servicios de gobierno digital en unidades académicas de nivel universitario

En el marco de las mejora de la calidad de los servicios institucionales, se inició una evaluación de la calidad de los servicios que brindan las unidades académicas (UA) con autonomía para definir sus procesos académicos y de gestión. [3]

El gobierno universitario está compuesto por docentes, no docentes, alumnos y graduados. Todos en su conjunto representan a la comunidad universitaria que desarrollan sus actividades en el marco de las reglamentaciones que dispone dicho gobierno. Para llevar a cabo el cumplimiento de las reglamentaciones, la universidad pone a disposición de su comunidad un conjunto de servicios. Actualmente, varios de estos servicios son brindados a través del uso TICs.

Se define el concepto de Gobierno Electrónico Universitario (EGOV-U), cómo "el uso de las TICs como herramienta para mejorar los procesos y los servicios prestados por una universidad a los miembros de su comunidad". [4]

Durante el 2017 se profundizó el modelo planteado y se está realizando la recolección de datos de las diferentes unidades académicas para validar el modelo.

4- Mejora de Procesos de gestión en el desarrollo de software

La dependencia creada por parte de la sociedad a las tecnologías de la información, la demanda generada por usuarios que reclaman servicios que den solución a sus necesidades y las

exigencias de un entorno cambiante, ha impulsado a un número significativo de compañías desarrolladoras de software a buscar soluciones enfocadas en mejorar organizacional y estratégicamente sus procesos para desarrollar productos de software de mayor calidad.

Los desarrollos de software tradicionales se han basado en un triángulo formado por el alcance, costo y duración de un proyecto. Este triángulo siempre ha supuesto que la calidad es inherente a los desarrollos, sin embargo, la misma ha resultado ser la variable que sufre en los proyectos.

El uso de las metodologías ágiles representa una alternativa para la mejora de procesos en el desarrollo de sistemas de software. Están centradas en el factor humano y en el producto de software, valorizando la relación con el cliente y el desarrollo incremental. Estas metodologías ofrecen entregas frecuentes de software funcional, permitiendo cambios de requerimientos y participación directa del cliente en el desarrollo.

Con el objetivo de asistir al equipo de desarrollo en el seguimiento del proyecto y almacenamiento de la documentación se analiza la posibilidad de brindar una herramienta que combine las funcionalidades que el equipo considere necesarias y las mismas se actualicen de forma automática en función de las modificaciones de los archivos del repositorio. Por ejemplo combinando las tareas de un tablero de Trello con los archivos en GitHub [5]–[7]

Por otro lado, en el camino de obtener mejor calidad en los procesos de desarrollo, se propone investigar la compatibilidad de la utilización de metodologías ágiles junto con los estándares de calidad.

5- Calidad en productos.

En la actualidad, el número de empresas desarrolladoras de software ha

experimentado un fuerte crecimiento, juntamente con el incremento de la demanda de productos del sector. Para este tipo de empresas, la calidad del software tiene un papel fundamental y las actividades relacionadas con la calidad de software junto con su evaluación están cobrando cada vez más importancia.

En este sentido, se ha estudiado la familia ISO/IEC 25000 conocida como SQuaRE (Software Product Quality Requirements and Evaluation), que posibilita la certificación de los productos de software. [8]

Dentro de esta familia de normas se destaca la norma ISO/IEC 25040, la cual define un proceso para llevar a cabo la evaluación de un producto de software. Dicha evaluación consta de una serie de pasos a seguir en los cuales se analizan diferentes aspectos y puntos de vista del producto. Uno de estos pasos consiste en seleccionar las características a evaluar, las cuales se encuentran definidas en el modelo de calidad propuesto por la norma ISO/IEC 25010.[9], [10]

A la hora de realizar una evaluación, el principal objetivo es obtener un documento detallado en el cual se posea la información necesaria para determinar el cumplimiento del propósito de la evaluación.

Para ello se desarrolló una herramienta denominada SEP (Sistema de Evaluación de Producto), la cual puede ser utilizada como guía por personas no experimentadas en el área de calidad de software, para realizar la evaluación de sus productos. Dicha herramienta no se utiliza con el fin de obtener una certificación, sino como un plan de mejora.

Mediante la utilización del enfoque GQM, se generó una alternativa para la medición propuesta por la norma ISO/IEC 25023, definiendo un conjunto de preguntas

cuyas respuestas combinadas de forma lógica permiten obtener una métrica aplicable a las características de la ISO/IEC 25010.[11][12]

Líneas de investigación y desarrollo

- Análisis y estudio de normas y modelos orientados a la Calidad del Producto, como por ejemplo IRAM - ISO/IEC 25000
- Evaluación de procesos en organismos públicos y privados según los requisitos de ISO 9001.
- Desarrollo de Voto electrónico presencial y Voto por Internet. Arquitecturas adaptadas a la legislación vigente.
- Análisis, discusión y estudio de normas de calidad relacionadas con certificación de servicios gubernamentales.
- Análisis, discusión y estudio de mejoras de proceso en el desarrollo de software combinando herramientas de gestión de proyectos.
- Evaluación de madurez de los servicios de gobierno digital de una unidad académica y de organismos públicos gubernamentales.

Resultados obtenidos / esperados

- Se avanza en la recolección de datos de las diferentes unidades académicas para validar el modelo de evaluación propuesto.
- Se avanza en el análisis de los servicios públicos digitales de acción social en los Municipios de La Plata y Bahía Blanca.
- Re-Certificación de los procesos Pre-Ingreso a Distancia de la Facultad de Informática y Concursos Docentes de la Facultad de Informática. Bajo la nueva versión de la normas IRAM-ISO 9001:2015.

- Análisis de nuevos procesos de la Facultad, con posibilidad de generar guías/instructivos para facilitar la prestación de los servicios y analizar la posibilidad de su certificación ser certificados.

- Se avanza en una herramienta para la asistir a la gestión de proyecto, generando documentación automática en función de los avances del desarrollo, combinando herramientas como Trello y GitHub
- Capacitación y desarrollo de los documentos básicos de gestión de la calidad de productos.
- Se implementó en una herramienta para asistir en la evaluación del producto de software utilizando métricas GQM
- Acciones de consultorías y asesoramiento en organismos públicos y privados
- Se avanzó en la tesis de doctorado “Modelo de madurez de los servicios de gobierno electrónico en el ámbito universitario”.
- Aplicación del Voto electrónico presencial en elecciones universitarias y Voto por Internet en distintos organismos.

Formación de recursos humanos

- Capacitación de los miembros del proyecto a través de diversos cursos del Instituto Argentino de Normalización y Certificación (IRAM).
- Se desarrollan tesis de doctorado y tesinas de grado en el área.
- Los integrantes de esta línea de Investigación participan en el dictado de asignaturas/cursos de grado/postgrado en la Facultad de Informática de la UNLP y en otras universidades del país. En particular, en la UNLP, se dicta la asignatura “Calidad de Sistemas de Software”.

Referencias

- [1] E. Estévez and T. Janowski, "Gobierno Digital, Ciudadanos y Ciudades Inteligentes.," 2016.
- [2] ISO, *IRAM - ISO 9001:2008 "Quality management systems -- Requirements."* 2008.
- [3] A. Pasini and P. Pesado, "Quality Model for e-Government Processes at the University Level: a Literature Review," *Proc. 9th Int. Conf. Theory Pract. Electron. Gov.*, pp. 436–439, 2016.
- [4] A. Pasini, E. Estévez, P. Pesado, and M. Boracchia, "Una metodología para evaluar la madurez de servicios universitarios," *Proc. Congr. XXII Congr. Argentino Ciencias la Comput. Congr. Argentino Ciencias la Comput.*, pp. 636–646, 2016.
- [5] C. W. Tom Preston-Werner PJ Hyett, "GitHub," 2008. 2013.
- [6] T. Inc, "About Trello," <https://trello.com/about>, 2016.
- [7] Atlassian, "JIRA Software," *Atlassian Foundation*, 2016. [Online]. Available: <https://www.atlassian.com/software/jira>.
- [8] ISO, "ISO/IEC 25000:2014 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Guide to SQuaREtle," 2014.
- [9] ISO, "ISO/IEC 25040:2011 □ Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Evaluation process," 2011.
- [10] ISO, "ISO/IEC 25010:2011 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models," 2011.
- [11] D. St-Louis and W. Suryn, "Enhancing ISO/IEC 25021 quality measure elements for wider application within ISO 25000 series," 2012, pp. 3120–3125.
- [12] J. Calabrese, R. Muñoz, A. Pasini, S. Esponda, and M. Boracchia, "Asistente para la evaluación de características de calidad de producto de software propuestas por ISO / IEC 25010 basado en métricas definidas usando el enfoque GQM," pp. 660–671, 2017.

Metodologías, técnicas y herramientas de ingeniería de software en escenarios híbridos.

**Rodolfo Bertone, Pablo Thomas, Ariel Pasini, Luciano Marrero, Eduardo Ibáñez,
Alejandra Rípodas, Verónica Aguirre, Verena Olsowy, Magalí Capecchi,
Fernando Tesone, Patricia Pesado**

Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – Universidad Nacional de La Plata

50 y 120 La Plata Buenos Aires

Centro Asociado CIC

{pbertone, pthomas, apasini, lmarrero, eibanez, aripodas, vaguirre, volsowy, mcapecchi, ftesone,
ppesado}@lidi.info.unlp.edu.ar

Resumen

Se presenta una línea de investigación y desarrollo que tiene por objeto estudiar y caracterizar los procesos relacionados con el diseño y desarrollo de sistemas de software para escenarios híbridos. Se tiene como objetivo principal generar metodologías y prácticas de Ingeniería de Software considerando las características de estos escenarios, así como realizar un análisis cualitativo y cuantitativo de aspectos de usabilidad que caracterizan los sistemas generados vinculados a contextos donde afectan la movilidad, la localización, las redes y los dispositivos involucrados.

Palabras claves: Escenarios Híbridos - Computación Ubicua - Usabilidad

Contexto

Esta línea de Investigación forma parte del proyecto (2018-2021) “*Metodologías, técnicas y herramientas de ingeniería de software en escenarios híbridos. Mejora de proceso.*”, en particular del subproyecto “*Ingeniería de Software para escenarios híbridos*”, del Instituto de Investigación en Informática LIDI (III-LIDI) de la Facultad de Informática UNLP, acreditado por el Ministerio de Educación de la Nación.

Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa en proyectos financiados por el Ministerio de Ciencia y Tecnología de España y la AECID.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Introducción

La creciente disponibilidad y acceso a dispositivos móviles, a redes inalámbricas, a sensores, sumado al uso del GPS y otras tecnologías que actualmente se encuentran al alcance de las personas, han marcado la profundización de nuevos entramados entre mundo físico y virtual y la puesta en práctica de los principios de la computación ubicua. Al mismo tiempo, las posibilidades de la web y sus desarrollos continúan siendo tema de investigación y actualidad.

Estos escenarios son reconocidos como híbridos, en ellos se aumenta computacionalmente a diferentes objetos del entorno y es posible acceder a la información digital, ya no sólo desde una PC de escritorio, sino desde dispositivos con variadas características. Los escenarios mencionados

modifican el concepto de ciudad (dando lugar a ciudades inteligentes), se integran en los escenarios educativos extendiendo las posibilidades del aula física, abren posibilidades para las actividades cotidianas y también llaman a la investigación y reflexión en el ámbito de las Ciencias de la Computación. Impulsan, además, innovaciones en las formas de gestión, como en el caso de los sistemas vinculados a la gobernanza electrónica. En este sentido, se hace necesario planificar metodologías y diseñar tecnologías que aprovechen estos escenarios híbridos complejos y con nuevas oportunidades para los usuarios [18].

Estos desafíos deben ser revisados y abordados desde la Ingeniería de Software, la atención de la calidad de los procesos y sistemas que involucran, los paradigmas de interacción persona-computadora para que permitan integrar con mayor naturalidad el uso de estos sistemas, y el desarrollo de interfaces que se adapten al contexto [2], [3], [5] [8] [14].

El desarrollo de software para escenarios híbridos requiere modificación de las prácticas actuales de Ingeniería de Software, y de aspectos de gestión de proyectos de desarrollo de software. Esas prácticas, ligadas a los modelos de ciclo de vida de desarrollo de software, ya sean ágiles o tradicionales, comprenden actividades de Ingeniería de Requerimientos, Diseño, Codificación y Prueba [1], [2], [3], [4] [6] [7] [9].

El proyecto *“Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso”*, 2018-2021 del Programa de Incentivos, propone profundizar en las investigaciones que ya se vienen realizando en estos temas en el III-LIDI y extender la mirada a estos nuevos desafíos y cambios que están en gestación. Para ello se organiza en tres subproyectos que permiten atender de manera ordenada el objetivo general propuesto: *“SP1 - Ingeniería de Software para escenarios híbridos”*, *SP2 - Gobernanza Digital. Mejora de Procesos.”* Y *“SP3 - Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos”*

En este artículo detallaremos el subproyecto *“SP1 -Ingeniería de Software para escenarios híbridos”*, se orienta a la investigación de metodologías y técnicas de la Ingeniería de Software, con énfasis en los escenarios híbridos, y bajo la consideración de las características de movilidad de las personas, uso de sistemas y aplicaciones desde diferentes dispositivos y puntos de acceso, y la necesidad de personalización y adaptación acorde al contexto y al estado de cada usuario. De manera particular en este subproyecto, se abordan los requerimientos para el diseño de aplicaciones para dispositivos móviles [12].

El segundo y tercer subproyecto, se orientan a la aplicación e integración de las metodologías, técnicas y herramientas que se investigan como parte del proyecto general, en dos áreas de aplicación concretas que también se ven impactadas por estos escenarios híbridos.

Las bases en que se asientan los métodos tradicionales de desarrollo de software son erosionadas a medida que los ingenieros de hardware crean distintos dispositivos con capacidad de cómputo (computadoras de escritorio, tablets, teléfonos inteligentes, entre otros), y/o mejoran las capacidades de los dispositivos existentes.

Por otra parte, actualmente los sistemas de software se utilizan en diferentes ámbitos. Una operación puede iniciarse desde un dispositivo y finalizarse desde otro. Por ejemplo, se puede iniciar una compra online agregando ítems a un carrito de compras desde un celular y finalizarla efectivizando el pago desde en una notebook; o iniciar la reproducción de una película desde un Smart-TV y continuar luego desde una tablet [11], [17].

También avanza la concepción de los sistemas ubicuos, incorporando la posición geográfica a los dispositivos, permitiendo procesar la información en función de su ubicación.

En el desarrollo de este tipo de aplicaciones, es vital el uso de técnicas para compartir información como los *Repositorios*, por

ejemplo GIT. El uso de estos repositorios permite realizar control de versiones distribuido, trabajando en modo off line o en modo online, con la facilidad de disponer herramientas específicas para la resolución de conflictos entre versiones [11], [17].

Un punto de interés de las aplicaciones constituye la sincronización de Bases de Datos Móviles con Bases de Datos Centralizadas o en la Nube. En los últimos años, la aparición de diferentes entornos para el funcionamiento de aplicaciones, ha generado la necesidad de operar tanto en modo online como off line, debiéndose sincronizar de información entre los datos de dichos entornos [10], [16].

Finalmente, otro caso distintivo es la computación móvil. La misma puede definirse como un entorno de cómputo con movilidad física. Un usuario debe ser capaz de acceder a datos, información u otros objetos lógicos desde cualquier dispositivo en cualquier red, al mismo tiempo que cambia su locación geográfica [17].

El desarrollo de aplicaciones para escenarios híbridos y con características de movilidad plantea, por lo tanto, nuevos desafíos originados en las características únicas de esta actividad. La necesidad de tratar con diversos estándares, protocolos y tecnologías de red; las capacidades limitadas, aunque en constante crecimiento de los dispositivos; las restricciones de tiempo impuestas por un mercado altamente dinámico, y la existencia de distintas plataformas de hardware y software son sólo algunas de las dificultades a las que se enfrentan los ingenieros de software en esta área [13], [19].

Todas las particularidades previamente mencionadas hacen que el desarrollo de software para dispositivos móviles, y para escenarios híbridos difiera considerablemente del tradicional. Esto conduce a nuevas prácticas y metodologías que promueven el crecimiento de la Ingeniería de Software como disciplina. [15].

Líneas de Investigación y Desarrollo

- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para escenarios híbridos.
- Sistemas Ubicuos
- Técnicas, atributos y métricas de usabilidad en sistemas para escenarios híbridos
- Metodologías para la interoperabilidad de aplicaciones móviles y sistemas web.
- Repositorios GIT.
- Trabajo colaborativo en Proyectos de Software
- Versionado Semántico de Software

Resultados esperados/obtenidos

Los resultados esperados/obtenidos se pueden resumir en:

- Avanzar en la capacitación continua de los miembros de la línea de investigación.
- Revisión, comparación y adaptación de metodologías y herramientas de la Ingeniería de Software para escenarios híbridos.
- Definición de técnicas, atributos y métricas para el análisis de usabilidad de sistemas en escenarios híbridos.
- Definición de procesos de Gestión de Incidencias utilizando repositorios GIT.
- Análisis de metodologías para la interoperabilidad de sistemas web y aplicaciones móviles.

Algunas transferencias realizadas por el Instituto fuertemente relacionadas con este proyecto son:

- Sistema de Alerta temprana. Sistema WEB de Alerta Temprana para la prevención de inundaciones, en el marco del proyecto PIO CONICET-UNLP que muestra los datos recibidos de un conjunto de pluviómetros. Permite monitorear no sólo las precipitaciones sino el crecimiento de cada cuenca hidrológica de la región.
- Sistema de Gestión Integral para el CONICET La Plata. Abarca las áreas

de Recursos Humanos; Comercio Exterior, Compras y Patrimonio; Administración Contable; Tesorería. Además dispone de funcionalidad utilizable en forma distribuida por todos los Centros de Investigación de la región dependiente del CCT La Plata.

- Sistema de Gestión de trazabilidad de semillas para la siembra. Consiste en la reingeniería del Sistema de control de calidad de semillas para la siembra en la Argentina operado por el INASE (Instituto Nacional de La Semilla).
- Sistema de Gestión integral de carreras de Postgrado. Se realizó el análisis y desarrollo de un Sistema WEB para gestión de carreras de Postgrado de la Facultad de Ciencias Económicas. Universidad Nacional de La Plata.
- Consultoría para la Secretaría de Comunicaciones de la Nación. Análisis de la viabilidad técnica de diferentes propuestas empresariales del programa Servicio Universal.
- Sistemas WEB para relevamiento de ofertas tecnológicas. Análisis y desarrollo de varios sistemas para distintos organismos (Provincia, UNLP, Universidades Nacionales).
- Sistema Inventario para Administrar Sustancias Utilizadas en las Unidades Ejecutoras y Cátedras de La Facultad de Ciencias Exactas de la UNLP (Abipon).

Formación de Recursos Humanos

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con el proyecto. Además participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

Referencias

[1] G. Coulouris. Distributed Systems – Concepts and Design. Addison-Wesley. 1994.

[2] R. Pressman. Ingeniería de Software: Un Enfoque Práctico. McGraw-Hill. 2002

[3] Pleegeer. Ingeniería de Software: Teoría y Práctica. Prentice-Hall. 2002

[4] Stephen Kan. Metrics and Models in Software Quality Engineering (2nd Edition). Addison Wesley. 2003

[5] Offutt J., “Quality Attributes of Web Software Applications”. IEEE Software: Special, Issue on Software Engineering of Internet Software 19 (2):25-32, Marzo / Abril 2002.

[6] Wu, Y. y Offutt, J. “Modeling and testing web-based Applications”. <https://citeseer.ist.psu.edu/551504.html>: 1-12, Julio 2004

[7] Ingeniería de Software, Ian Sommerville, Pearson, Addison Wesley, 2006

[8] Software Engineering Institute. Disponible en <http://www.sei.cmu.edu/risk/>

[9] A spiral model of software development and enhancement. B. Boehm, IEEE Computer, 21(5), 61-72. 1988

[10] Amja, A.M., Obaid, A., Seguin, N. (2011). *A Distributed Mobile Database Architecture*. 2011 IEEE Asia -Pacific Services Computing Conference.

[11] Cosentino, V., Canovas Izquierdo, J.L., Cabot, J. (2015). *Assessing the Bus Factor of Git Repositories*, SANER, Montréal, Canada. IEEE.

[12] Delia, L., Galdamez, N., Corbalan, L., Thomas, P., Pesado, P. (2015). *Multi-Platform Mobile Application Development Analysis*, IEEE Ninth International Conference on Research Challenges in Information Science, May 2015, Athens, Greece, ISBN 978-1-4673-6630-4

[13] IRAM-ISO/IEC 14598 (s.f.). *Evaluación del Producto de Software*

[14] Ishii, H. & Ullmer, B. (1993). *Tangible bits: towards seamless interfaces between people, bits and atoms*. In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'97), pp. 234-241

[15] ISO/IEC 9126. (s.f.). *Software engineering -- Product quality*

[16] Li, S., Gao, J. (2015). *Moving from Mobile Databases to Mobile Cloud Data Services*. 2015 3rd IEEE International

Conference on Mobile Cloud Computing, Services, and Engineering.

[17] Ramadani, J., Wagner, S. (2016). *Which Change Sets in Git Repositories Are Related?* 2016 IEEE International Conference on Software Quality, Reliability and Security, Viena, Austria.

[18] Rösler, F., Nitze, A., Schmietendorf, A. (2015). *Towards a Mobile Application Performance Benchmark*. ICIW 2014: The Ninth International Conference on Internet and Web Applications and Services, At Paris, France.

[19] Wasserman, A. (2011). *Software Engineering Issues for Mobile Application Development*, 2° Workshop on Software Engineering for Mobile Application Development, MobiCASE '11, Carnegie Mellon Silicon Valley, Santa Monica, California, USA, October 2011

Coordinación de dispositivos ubicuos: una solución basada en SOA y coreografías

Germán Montejano^{1,2}; Oscar Testa²; Rubén Pizarro²; Darío Segovia²; Oscar Dieste³; Efraín R. Fonseca C.⁴

¹Facultad de Ciencias Físico-Matemáticas y Naturales
Universidad Nacional de San Luis
Ejército de los Andes 950 – (5700) San Luis – San Luis – Argentina
Tel.: +54-2652-424027 – Int. 251
gmonte@unsl.edu.ar – web: <http://www.unsl.edu.ar>

²Facultad de Ciencias Exactas y Naturales
Universidad Nacional de La Pampa
Av. Uruguay 151 – (6300) Santa Rosa – La Pampa – Argentina
Tel.: +54-2954-425166 – Int. 28
[otesta, ruben]@exactas.unlpam.edu.ar

³Facultad de Informática
Universidad Politécnica de Madrid
Campus Montegancedo – (28660) Boadilla del Monte – Madrid – España
Tel.: +34 91 336 5011
odieste@fi.upm.es

⁴Departamento de Ciencias de la Computación
Universidad de las Fuerzas Armadas ESPE
erfonseca@espe.edu.ec

Resumen

Actualmente nos encontramos involucrados en ambientes donde los dispositivos ubicuos forman parte de nuestra vida cotidiana y de nuestras tareas diarias. De forma permanente estamos interactuando con dichos dispositivos y más aún, con los servicios que ellos nos brindan. En casi todos los casos, los dispositivos ubicuos no proporcionan servicios de forma aislada, sino que deben cooperar con otros dispositivos. Actualmente, los mecanismos de cooperación disponibles son fundamentalmente de tipo propietario, y las pocas propuestas provenientes de la academia no han tenido apenas impacto en la práctica.

Nuestra propuesta es adaptar y aplicar las especificaciones de coreografías actualmente existentes en SOA para la coordinación de servicios proporcionados por dispositivos ubicuos. Para poder cumplir con el objetivo de tesis planteado se utilizará como metodología de investigación design science, ya que es la que mejor se adapta a la naturaleza del problema, planteando como uno de sus lineamientos la contrucción de artefactos y su posterior evaluación. En nuestro caso, el

artefacto a construir sería: una especificación de coreografías adaptado a los dispositivos ubicuos y un framework que implementa dicha especificación al nivel de prueba de concepto.

Palabras clave: sistemas ubicuos, SOA, servicios, composición de servicios, coreografías.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software, Conceptos, Métodos y Herramientas en un Contexto de “Ingeniería de Software en Evolución” – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis y en el Proyecto de Investigación: Ingeniería de Software: composición de servicios en ambientes ubicuos – Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa. Las líneas aquí presentadas actualmente forman parte de las bases de un anteproyecto de tesis doctoral y de una tesis de maestría, ambas en ingeniería de software por la Universidad Nacional de San Luis.

Introducción

En la actualidad, las actividades cotidianas del hombre se han hecho dependientes de una gran cantidad de dispositivos electrónicos tales como: ordenadores personales, ordenadores portátiles, teléfonos móviles, PDAs, tabletas, sensores de muchas y diversas utilidades, entre otros; los cuales logran comunicarse entre sí gracias a diversos protocolos de comunicación inalámbrica, redes de celulares, redes de área local (LAN), redes de área extensa (WAN), Bluetooth, etc. Estamos en la presencia de nuevos dispositivos de comunicación, lo que conlleva un nuevo escenario social, donde la interacción permanente con estos elementos es ineludible. Por ejemplo, la cantidad de móviles existentes en el mercado se aproxima a la cantidad de habitantes mundiales, según un informe de la Unión Internacional de Comunicaciones. En este informe, se estima que hasta finales del año 2014 hubo casi 7.000 millones de suscripciones de telefonía celular, lo que corresponde al 96 % de la población global; es decir, la cantidad de usuarios de telefonía móvil se acerca al número de personas que viven en el planeta [1].

Los avances de las comunicaciones entre dispositivos ha permitido que estos sean generadores y consumidores de servicios al mismo tiempo, es decir, de acuerdo a las capacidades del dispositivo puede no solo obtener, sino también ofrecer a otros equipos sus funciones y así cooperar entre ellos. La tendencia actual es hacia los ambientes ubicuos, los cuales se caracterizan por estar poblados de numerosos dispositivos que, gracias a la integración extrema de los elementos electrónicos, son invisibles al usuario y están en permanente rastreo de la actividad humana [2].

Dispositivos ubicuos son todos aquellos dispositivos que pueden existir en todas partes, es decir, son dispositivos electrónicos que tienen capacidad de procesamiento y comunicación y pueden ser encontrados en lugares diversos de la vida cotidiana.

La computación ubicua es un desarrollo tecnológico que intenta que las computadoras no se perciban en el entorno como objetos diferenciados, y que la utilización por parte de los seres humanos sea lo más transparente y cómoda posible, facilitando de esta manera la integración en la vida cotidiana. Desde hace varios años los dispositivos ubicuos han ganado importancia y presencia en la vida cotidiana de las personas, debido principalmente a que: poseen distintos tipos de sensores (posicionamiento, proximidad, luminosidad, temperatura, etc.), facilitan la conectividad incluso en áreas con poca señal o acceso a las redes, permiten la convergencia tecnológica (computo, medios, telefonía, etc) y brindan acceso a servicios de distinta índole (mapas, ayudas, etc).

Por composición entendemos la forma en que se pueden combinar o enlazar un número indeterminado de dispositivos para llevar adelante una tarea determinada. En ambientes ubicuos, la composición de dispositivos, presenta nuevos desafíos tales como: la heterogeneidad (ya sea por la diversidad de dispositivos involucrados, como por la presencia de dispositivos de varios fabricantes), las contingencias de los dispositivos y la personalización de los mismos (por ej. provisión de servicios de acuerdo a las preferencias del usuario). Dado que los dispositivos en donde los servicios son ejecutados poseen limitaciones de recursos (ej. poca memoria y batería), se deben hacer consideraciones especiales respecto a la eficiencia y rendimiento de la composición de servicios [3].

La composición en este tipo de ambientes implica que los dispositivos deben dialogar entre ellos para poder compartir los servicios que ofrecen con la finalidad de obtener un servicio con valor agregado, o bien para abordar la solución de una problemática particular, como podría ser la seguridad de un hogar, o la seguridad vial, por mencionar algunos ejemplos.

Si bien hoy en día podemos decir que distintos sensores o dispositivos se pueden comunicar entre ellos, compartiendo de alguna manera sus servicios, generalmente lo realizan a partir de protocolos propietarios y sin seguir definiciones estándares, provocando que otros componentes de otros proveedores (o incluso de los mismos) no puedan ser utilizados. Esto obviamente representa una importante limitación en la composición de dispositivos ubicuos. Adicionalmente la composición de dispositivos ubicuos presenta un nuevo desafío. Los mecanismos de composición en ambientes masivos, necesitan hacer frente a las distintas contingencias que pueden ocurrir con estos dispositivos. Los dispositivos ubicuos tienen distintas limitantes como son la cantidad de memoria disponible, la durabilidad de la batería, la disponibilidad de acuerdo a la red del lugar donde se encuentre en un momento determinado. Todas estas variantes hacen que la composición de dispositivos¹ ubicuos se transforme en un área de investigación muy importante donde los avances no han sido claros al día de hoy[3].

Líneas de Investigación y Desarrollo

La computación orientada a servicios, y en particular los servicios web en ambiente de internet, proporcionan mecanismos para la composición de servicios. Dichos mecanismos, como las orquestaciones, son aspectos bien conocidos de la computación orientada a servicios que permiten construir sistemas de negocio complejos y aplicaciones a partir de una gran cantidad de servicios heterogéneos, simples y distribuidos. Podría pensarse que son aplicables a ambientes ubicuos. Sin embargo,

¹ Si bien los autores se refieren a la composición de servicios, se hace dentro de un contexto de dispositivos ubicuos, lo cual a los fines de este trabajo se puede interpretar como composición de dispositivos, haciendo que la terminología para este caso particular sea más adecuada.

en contextos como puede ser la Internet de las Cosas (IoT) donde los servicios son dinámicos, móviles, menos fiables y dependientes del dispositivo, los mecanismos de composición establecidos para servicios web no es directamente aplicable [8].

Adicionalmente la composición de múltiples dispositivos ubicuos presenta nuevos desafíos que no son compatibles con la composición de servicios web. En particular, los mecanismos de composición en ambientes masivos como lo es el de dispositivos móviles, necesita hacer frente las distintas contingencias que pueden ocurrir con estos elementos, así como también contemplar la heterogeneidad de los mismos.

Estos dispositivos tienen distintas limitantes como son la cantidad de memoria disponible, la durabilidad de la batería, la disponibilidad de acuerdo a la red del lugar donde se encuentre en un momento determinado. En ambientes ubicuos, la disponibilidad y confiabilidad de los dispositivos no puede ser garantizada. Todas estas dificultades hacen que la composición de dispositivos se transforme en un área de investigación muy importante donde los avances no han sido claros al día de hoy[3].

Finalmente existen distintos proyectos en la actualidad donde se intenta integrar sensores y dispositivos ubicuos a la vida cotidiana. Específicamente podemos mencionar la domótica, donde varios dispositivos y sensores deben actuar en coordinación para prevenir un incidente de seguridad (ya sea por robo o por incendio) en nuestros hogares. Sin embargo, existen áreas de aplicación más relevantes.

En la industria, existe lo que se llama Industria 4.0 [9], donde lo que se intenta es integrar dentro de una planta fabril la intercomunicación de todos los dispositivos que componen la cadena de producción con el fin de que coordinen entre ellos las tareas a realizar en base a los tiempos a cumplir, stocks disponibles, demanda en línea de los productos, etc. Otra área donde los dispositivos ubicuos están ganando importancia es la automotriz, donde los esfuerzos se enfocan en que distintos sensores monitoreen funciones vitales del

conductor (como es el caso de presión arterial, pulsaciones, etc) y en caso de que detecten anomalías actúen en conjunto con otros dispositivos del vehículo para evitar accidentes.

Es claro que en este punto se hace necesaria una mayor investigación y desarrollo de tecnologías que permitan solucionar en todo o en parte estos desafíos planteados, haciendo foco en la composición de distintos dispositivos de una manera abierta y estándar.

Los mecanismos de composición en ambientes ubicuos como los dispositivos móviles, necesitan hacer frente a las distintas contingencias que pueden ocurrir con estos dispositivos, así como también contemplar la heterogeneidad de los mismos. La heterogeneidad no sólo se refiere a la existencia de dispositivos de distintos modelos, sistemas operativos y fabricantes, sino también a los mecanismos de comunicación e interacción que poseen los mismos, en algunos casos propietarios, provocando que la interacción y coordinación entre ellos representa un desafío de enorme magnitud. Estos dispositivos tienen, a su vez, limitantes adicionales como son la cantidad de memoria disponible, la durabilidad de la batería o la conectividad de acuerdo a la red del lugar donde se encuentre en un momento determinado. Todas estas dificultades hacen que la composición de servicios incluyendo dispositivos móviles se transforme en un área de investigación muy importante donde los avances no han sido claros al día de hoy[3].

A medida que los dispositivos ubicuos son menos potentes (ej: cámaras de seguridad, sensores, etiquetas RFID, etc), la disponibilidad y confiabilidad de los mismos no puede ser garantizada. En este tipo de ambientes, mecanismos automáticos y dinámicos son necesarios para la composición de dispositivos, ya que de esta forma se puede compensar la falta de disponibilidad de un dispositivo en un momento determinado[8].

Resultados y Objetivos

Por lo expuesto, vemos que existe un campo de trabajo importante en el desarrollo de composición de servicios en ambientes ubicuos, más precisamente en la coreografía de servicios, la cual no es abordada en los estudios previos de la materia.

Por todo esto, nuestra propuesta es poder adaptar y aplicar las especificaciones actualmente existentes en SOA para la coordinación de servicios disponibles en ambientes pervasivos a través de la utilización de dispositivos ubicuos, más concretamente, el objetivo de esta investigación es:

- Definir un mecanismo de coordinación de dispositivos ubicuos que garantice su interoperabilidad independientemente del modelo y fabricante del mismo; utilizando los estándares de SOA y de coreografías para la composición de servicios.

Debemos destacar que la aplicación de los conceptos de SOA a dispositivos ubicuos no consiste en una mera traslación de los conceptos de un ambiente a otro, sino que será necesario para ello extender las especificaciones de SOA existentes de modo que se adapten a las circunstancias particulares de los sistemas ubicuos. Asimismo se deberá mantener total compatibilidad con las especificaciones relacionadas a SOA y coreografía de servicios existentes.

En relación al estado actual de la investigación, podemos decir que hemos logrado realizar un framework de ejecución de coreografías especificadas a través de WSDL. Se han alcanzado a realizar 3 pruebas de concepto de simulación de un escenario que involucra dispositivos ubicuos. Para llevar adelante estas ejecuciones se han utilizado dispositivos de muy pequeña capacidad como son las placas Arduino Mega y Arduino Nano, donde las capacidades tanto de procesamiento como de almacenamiento son muy escasas.

Son varios los aportes que se han necesitado realizar al lenguaje de especificación de coreografías WS-CDL para ser adaptado para la ejecución de coreografías en ambientes ubicuos, en línea con el objetivo de tesis planteado. Hasta el momento podemos mencionar: la adaptación del concepto de paralelismo en coreografías a través de llamadas secuenciales debido a la falta de un sistema operativo que brinde la capacidad de multiprocesamiento en placas Arduino; la transformación de la descripción de la coreografía en lenguaje XML a un conjunto de vectores en memoria con la información necesaria para que se ejecute la coreografía desde los dispositivos con menor capacidad de memoria; la codificación de un motor de ejecución de coreografías para distintos dispositivos (con adaptaciones diversas para los que tienen menores capacidades). Sería necesario realizar además otros aportes para adaptar WS-CDL a algunas de las características de los sistemas ubicuos como pueden ser desapariciones por problemas de red, falta de batería u otros.

Formación de Recursos Humanos

Además de los resultados obtenidos/esperados en el punto 3, se espera como resultado en la formación de recursos humanos, la continuación de esta misma línea de proyecto como tesis doctoral de alguno(s) de los investigadores. También se espera lograr una mayor interrelación con la Universidad de Minas Gerais con la que se cuenta con un convenio con tal objetivo como parte de él. Se espera avanzar también en un convenio de colaboración con la Universidad Politécnica de Madrid para la aplicación de las metodologías aquí presentadas en los proyectos de Ingeniería de Software Empírica. Adicionalmente, se espera que otras tesis de Maestría, así como tesis de Licenciatura surjan a partir de los logros obtenidos en la presente línea de investigación.

Bibliografía

- [1] U. I. d. T. (UIT), "Unión Internacional de Telecomunicaciones." <https://www.itu.int/net/pressoffice/pressreleases/2014/23-es.aspx>, 10 2015.
- [2] M. Weiser, "Hot topics-ubiquitous computing," *Computer*, vol. 26, pp. 71–72, Oct 1993.
- [3] Q. Z. Sheng, X. Qiao, A. V. Vasilakos, C. Szabo, S. Bourne, and X. Xu, "Web services composition: A decade's overview," *Information Sciences*, vol. 280, no. 0, pp. 218–238, 2014.
- [4] M. Viroli, "On competitive self-composition in pervasive services," *Science of Computer Programming*, vol. 78, no. 5, pp. 556–568, 2013. Special section: Principles and Practice of Programming in Java 2009/2010 & Special section: Self-Organizing Coordination.
- [5] S. W. Loke, "Supporting ubiquitous sensor-cloudlets and context-cloudlets: Programming compositions of context-aware systems for mobile users," *Future Generation Computer Systems*, vol. 28, no. 4, pp. 619–632, 2012.
- [6] F. Palmieri, "Scalable service discovery in ubiquitous and pervasive computing architectures: A percolation-driven approach," *Future Generation Computer Systems*, vol. 29, no. 3, pp. 693–703, 2013. Special Section: Recent Developments in High Performance Computing and Security.
- [7] S. Najar, M. K. Pinheiro, and C. Souveyet, "A New Approach for Service Discovery and Prediction on Pervasive Information System," *Procedia Computer Science*, vol. 32, pp. 421–428, 2014. The 5th International Conference on Ambient Systems, Networks and Technologies (ANT-2014), the 4th International Conference on Sustainable Energy Information Technology (SEIT-2014).
- [8] G. Cassar, P. Barnaghi, W. Wang, S. De, and K. Moessner, "Composition of services in pervasive environments: A Divide and Conquer approach," in *Computers and Communications (ISCC), 2013 IEEE Symposium on*, pp. 000226–000232, July 2013.
- [9] Wikipedia, "Industria 4.0 — Wikipedia, La enciclopedia libre," 2016. [Internet; descargado 4-noviembre-2016].
- [10] H.-I. Yang, R. Bose, A. (Sumi) Helal, J. Xia, and C. Chang, "Fault-Resilient Pervasive Service Composition," in *Advanced Intelligent Environments (A. D. Kameas, V. Callagan, H. Hagra, M. Weber, and W. Minker, eds.)*, pp. 195–223, Springer US, 2009.

ENFOQUES DE OPTIMIZACIÓN MULTI-OBJETIVO BASADOS EN PREFERENCIAS EN LA INGENIERÍA DE SOFTWARE

Martín Arrúa, Luciano Bracco, Giovanni Daián Rottoli, Esteban Schab, Adrián Tournoud, Carlos Casanova, Anabella De Battista

Grupo de Investigación sobre Inteligencia Computacional e Ingeniería de Software, Departamento Ingeniería en Sistemas de Información, Fac. Reg. Concepción del Uruguay, Universidad Tecnológica Nacional
Entre Ríos, Argentina

{arruam, braccol, rottolig, schabe, tournoura, casanovac, debattistaa}@frcu.utn.edu.ar

RESUMEN

La Ingeniería de Software Basada en Búsqueda (ISBB) estudia la aplicación de técnicas de optimización metaheurística a problemas de la Ingeniería de Software (IS). Una vez que una tarea de la IS se enmarca en un problema de búsqueda existen multitud de algoritmos que pueden aplicarse para resolver ese problema. La mayoría del trabajo existente trata a los problemas de la IS desde un punto de vista mono-objetivo. Sin embargo, muchos de estos problemas poseen múltiples objetivos en conflicto que deben ser optimizados. El número de objetivos a considerar es, en general, alto (esto es, más de tres objetivos). Si bien la comunidad científica ha propuesto varios enfoques de solución para atacar la optimización multi-objetivo, muchos de estos enfoques nos se han aplicado aún en la ISBB. Uno de estos enfoques es el llamado “basado en preferencias”, el cual permite incorporar las preferencias entre los objetivos del tomador de decisiones, restringiendo el frente Pareto-óptimo a una zona de interés específica, facilitando de esta manera la tarea de tomar una decisión.

Palabras clave: Ingeniería de Software Basada en Búsqueda, Optimización Multi-objetivo, Metaheurísticas, Hiperheurísticas, Preferencias.

CONTEXTO

El presente trabajo se desarrolla en el ámbito del Grupo de Investigación sobre Inteligencia Computacional e Ingeniería de Software (GIICIS), perteneciente al Departamento Ingeniería en Sistemas de Información de la

Universidad Tecnológica Nacional, Facultad Regional Concepción del Uruguay.

1. INTRODUCCIÓN

Muchos de los problemas asociados con el manejo de la complejidad del proceso de desarrollo de software han probado ser resistentes a soluciones analíticas convencionales. El proceso de desarrollo de software y sus productos tienden a estar caracterizados por un gran número de restricciones y objetivos interrelacionados que compiten entre sí. Algunos de ellos están claramente especificados, mientras que otros son notablemente vagos y se encuentran pobremente definidos. Cambios en un parámetro con frecuencia tienen un gran impacto sobre áreas relacionadas, haciendo que el balance de intereses sea muy difícil de lograr. Estas características producen que las soluciones “perfectas” sean, si no imposibles, al menos imprácticas.

Si sólo hubiera una única solución para un conjunto típico de restricciones de un problema de la Ingeniería de Software, entonces probablemente no sería llamada “ingeniería”. De la misma manera que ocurre con otras disciplinas, la Ingeniería de Software generalmente se ocupa de soluciones cuasi-óptimas, o aquellas que caen dentro de una tolerancia aceptable. Es así que los ingenieros de software se enfrentan a problemas que consisten, no en encontrar *la solución*, sino en construir una solución aceptable, o cuasi-óptima, partiendo de una gran cantidad de alternativas. Con frecuencia no está del todo claro cómo alcanzar una solución óptima, pero puede diseñarse una forma de evaluar y

comparar candidatos. Por caso, puede ser difícil conocer cómo alcanzar un diseño con alta cohesión y bajo acoplamiento, pero es relativamente sencillo decidir cuándo un diseño está más acoplado que otro.

La siguiente es una lista ilustrativa de problemas de optimización de la ingeniería de software:

- ¿Cuál es el mínimo conjunto de casos de test que cubren todas las ramas de un programa?
- ¿Cuál es la mejor manera de estructurar la arquitectura de este sistema?
- ¿Cuál es el conjunto de requerimientos con el mejor balance entre costo de desarrollo y satisfacción de los clientes?
- ¿Cuál es la mejor asignación de recursos para este proyecto de desarrollo de software?
- ¿Cuál es la mejor secuencia de pasos de refactorización a aplicar en este sistema?

Las propiedades descriptas sobre la Ingeniería de Software son precisamente los atributos que hacen a la aplicación de técnicas basadas en búsqueda tan atractiva. Las técnicas de búsqueda metaheurística son un conjunto de algoritmos genéricos adecuados para la búsqueda de soluciones óptimas o cuasi-óptimas en problemas con un enorme espacio de búsqueda multimodal [12]. La aplicación de técnicas de búsqueda a problemas de la ingeniería de software recibe en el estado del arte el nombre de Ingeniería de Software Basada en Búsqueda (*Search-Based Software Engineering*, SBSE) [1].

Los algoritmos metaheurísticos brindan a sus usuarios estrategias de búsqueda genéricas, a la vez que requieren, para ser aplicados a un problema en específico, las siguientes definiciones (según [1]): (a) la forma de representación de las soluciones del problema; (b) una función de evaluación, definida para tal representación, que mide la calidad de las soluciones, y (c) un conjunto de operadores para manipular esas soluciones.

Es de destacar que, para resolver un problema mediante metaheurísticas, no es necesario realizar un modelado algebraico analítico de la misma manera que se realiza, por caso, en los enfoques basados en programación matemática. La función de evaluación puede estar definida, por caso, mediante un algoritmo, o utilizando varios algoritmos para su cómputo. Esta es una ventaja muy importante, mayormente en lo que se refiere a expresividad de los modelos.

Según lo expuesto hasta el momento, los problemas requieren soluciones con ciertos atributos o propiedades que pueden ser expresados como funciones que mapean posibles soluciones a valores numéricos escalares. Un enfoque para la optimización comúnmente usado es tomar esos atributos como restricciones para determinar la factibilidad de una solución, mientras que uno de ellos es elegido como función objetivo, la cual determina el orden (total) de preferencia de las soluciones factibles. Este tipo de problemas de optimización se refieren como mono-objetivo.

Un enfoque más general es la optimización multi-objetivo, donde varios atributos se emplean como funciones objetivo y se usan para definir un orden de preferencia parcial de las soluciones factibles. Por caso, en el problema nombrado más arriba sobre encontrar el subconjunto de requerimientos, se desea minimizar el costo de desarrollo y maximizar la satisfacción de los clientes. Estos objetivos, al igual que en muchos otros problemas, son de cierta forma contradictorios y compiten entre sí, definiendo sobre el espacio de soluciones un orden parcial, la relación de dominancia, donde existen pares de soluciones que no son comparables a priori. Ante esta situación pueden tomarse básicamente dos estrategias para la resolución del problema de optimización multi-objetivo:

- (a) buscar el frente de Pareto completo, compuesto por todas las soluciones no dominadas y dejar al tomador de decisiones la tarea de buscar en ese frente la solución que mejor se adecúe a su criterio (por caso, mediante otro algoritmo), o bien

- (b) utilizar algún mecanismo (por caso, una nueva relación de orden total) que reduzca el frente de Pareto a una única solución o a una zona de interés.

El enfoque descrito en el punto (a) ha sido estudiado en trabajos como [2], [3]. Una desventaja importante de este tipo de enfoques es que la performance de los algoritmos baja rápidamente cuando aumenta el número de objetivos. Se estima que para una cantidad de 4 o más objetivos este tipo de enfoques presenta algunos inconvenientes difíciles de mitigar, como frentes Pareto-óptimos muy grandes, alta dimensionalidad que dificulta encontrar buenas direcciones en la búsqueda, e incapacidad de la relación de dominancia para distinguir entre vectores objetivo [4].

Por otra parte, un mecanismo del tipo de los descritos en el punto (b) es utilizar información concerniente a las preferencias del tomador de decisiones entre los objetivos, esto es, establecer la importancia relativa de los objetivos. Los algoritmos que incorporan las preferencias, intuición o emoción en el proceso de optimización reciben el nombre de algoritmos basados en preferencias (ABP). Tanto la resolución de problemas multi-objetivo como la utilización de preferencias en el contexto de la ISBB han sido identificados en *reviews* recientes como desafíos y áreas emergentes [5]–[7].

Además de las tres definiciones que ya se han descrito como requeridas para la aplicación de técnicas basadas en búsqueda, se requiere un cuarto mecanismo para aplicar en este contexto un ABP: una forma de incorporar la información del tomador de decisiones relativa a las preferencias [7]. A grandes rasgos pueden identificarse tres tipos principales según el momento del proceso de optimización donde se incorporan las preferencias: a priori (antes de comenzar el proceso), interactivos (durante) y a posteriori (después). De ninguna manera estos enfoques son mutuamente excluyentes, pudiendo ser combinados de diversas maneras. Ejemplos de ABP utilizados en la ISBB son los algoritmos genéticos, siendo estos los más utilizados por un amplio margen, en particular el Algoritmo Genético Interactivo (IGA, de sus siglas en inglés), utilizado en varios trabajos.

También pueden encontrarse aplicaciones con ACO (*Ant Colony Optimization*), NSGA-II (*Nondominated Sorting Genetic Algorithm*), VEGA (*Vector Evaluated Genetic Algorithm*), IBEA (*Indicator Based Evolutionary Algorithm*), HCM (*Hard C-Mean*) y DE (*Differential Evolution*).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

2.1. Estudio de ABPs en distintos problemas de la IS

La ISBB es una disciplina relativamente nueva. Si bien ha cobrado mucha relevancia en los últimos años [7], todavía escasean los estudios empíricos sobre la habilidad exhibida por distintos métodos para resolver los distintos problemas de optimización derivados de la IS. Los ABP no son la excepción a esta regla. De modo que el trabajo principal en esta línea es identificar los ABP presentes en el estado del arte y aplicarlos a aquellos problemas de la ISBB en los cuales no se tiene registro de haberse aplicado hasta el momento. Claramente, el objetivo es realizar comparaciones respecto de diversos criterios: tiempo, calidad, robustez, facilidad de uso, nivel de generalidad, etc. Una dimensión más del estudio es analizar cómo cambia el comportamiento de cada ABP según se realice la incorporación de la información relativa a las preferencias de manera a priori, interactiva o a posteriori. Ejemplos ABP en otros dominios que no han sido aplicados aún pueden encontrarse en [8]–[10].

2.2. Diseño y desarrollo de nuevos algoritmos

Los algoritmos metaheurísticos disponibles en el estado del arte pueden no ser adecuados para la resolución de ciertas clases de problemas. Por caso, la metaheurística PSO original (*Particle Swarm Optimization*) supone un espacio de búsqueda continuo, y codifica las soluciones en vectores. Esta representación no resulta adecuada para resolver problemas de optimización combinatoria, como TSP (*Traveling Salesman Problem*), donde las soluciones son permutaciones de un conjunto.

Es así que las metaheurísticas pueden reformularse para atacar distintos problemas. Para el caso de PSO se cuenta con el *template* llamado PSO Canónico descrito en [11], el cual constituye una generalización de PSO que puede especificarse para resolver distintos problemas. Es así que para problemas específicos de la ISBB se trabaja en lograr algoritmos específicos novedosos que mejoren a los disponibles en algún aspecto. Otro tipo de algoritmos específicos en los que se trabaja es en las hiperheurísticas, las cuales constituyen métodos de un nivel de abstracción superior que buscan combinar distintas estrategias heurísticas o metaheurísticas para la resolución de un problema o una clase de problemas. Este tipo de algoritmos aún no ha sido aplicado ampliamente en la ISBB.

2.3. Mecanismos de captación de preferencias

Como ya se dijo, los enfoques de optimización multi-objetivo basados en preferencias suponen que puede incluirse información del tomador de decisiones para guiar la búsqueda hacia regiones prometedoras del espacio de soluciones. Idear mecanismos adecuados para captar esta información no es una tarea trivial, y se trabaja en el estudio de los mecanismos disponibles, y el diseño e implementación de nuevos mecanismos para mejorar principalmente la usabilidad de los métodos. Los componentes analizables de estos mecanismos son principalmente tres: la cantidad de interacciones con el usuario, qué información se le solicita, y en qué momento del proceso de optimización [7].

3. RESULTADOS OBTENIDOS/ESPERADOS

Se espera obtener una plataforma de pruebas homogénea para realizar comparaciones entre distintos ABP, utilizados con distintos esquemas de incorporación de preferencias. Para ello no se descarta la codificación o recodificación de algunos de los algoritmos en un lenguaje de programación de alta performance, como C++.

También se espera proponer mejoras en distintos ABP existentes, así como la adecuación de aquellos que no se encuentran adaptados para los problemas de la ISBB basada en preferencias.

Finalmente, se espera lograr una herramienta amigable y flexible que permita captar la información del tomador de decisiones de diversas maneras.

4. FORMACIÓN DE RECURSOS HUMANOS

Las líneas de investigación presentadas constituyen las líneas fundacionales de un nuevo grupo de investigación dentro de la UTN-FRCU, el GIICIS. Un investigador se encuentra realizando su tesis de doctorado. Un investigador está desarrollando su trabajo de especialización. Además participan en el proyecto tres becarios alumnos de la carrera Ingeniería en Sistemas de Información que inician su formación en la investigación. Finalmente, uno de los investigadores finalizó su doctorado en el mes de febrero de 2018 en la temática de métodos hiperheurísticos.

5. BIBLIOGRAFÍA

- [1] M. Harman y B. F. Jones, «Search-based software engineering», *Inf. Softw. Technol.*, vol. 43, n.º 14, pp. 833-839, 2001.
- [2] A. S. Sayyad, T. Menzies, y H. Ammar, «On the value of user preferences in search-based software engineering: A case study in software product lines», *2013 35th Int. Conf. Softw. Eng.*, pp. 492-501, 2013.
- [3] S. Z. Qasim y M. A. Ismail, «Research problems in Search-Based Software Engineering for many-objective optimization», en *2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT)*, 2017, pp. 1-6.

- [4] M. W. Mkaouer, M. Kessentini, S. Bechikh, y D. R. Tauritz, «Preference-based multi-objective software modelling», en *2013 1st International Workshop on Combining Modelling and Search-Based Software Engineering, CMSBSE 2013 - Proceedings*, 2013, pp. 61-66.
- [5] M. Harman, S. A. Mansouri, y Y. Zhang, «Search Based Software Engineering: Trends, Techniques and Applications», *ACM Comput. Surv.*, vol. 45, n.º 1, p. 11, 2012.
- [6] M. Harman, S. Mansouri, y Y. Zhang, «Search Based Software Engineering: A Comprehensive Analysis and Review of Trends Techniques and Applications», p. 78, 2009.
- [7] T. N. Ferreira, S. R. Vergilio, y J. T. de Souza, «Incorporating user preferences in search-based software engineering: A systematic mapping study», *Inf. Softw. Technol.*, vol. 90, pp. 55-69, 2017.
- [8] R. Malek, «An agent-based hyper-heuristic approach to combinatorial optimization problems», *Proc. - 2010 IEEE Int. Conf. Intell. Comput. Intell. Syst. ICIS 2010*, vol. 3, pp. 428-434, 2010.
- [9] G. Schweickardt, J. M. G. Alvarez, y C. Casanova, «Metaheuristics approaches to solve combinatorial optimization problems in distribution power systems. An application to Phase Balancing in low voltage three-phase networks», *Int. J. Electr. Power Energy Syst.*, vol. 76, pp. 1-10, 2016.
- [10] C. Casanova y G. Schweickardt, «Análisis de las habilidades de Metaheurísticas X-PSO multi-objetivo mediante Indicadores de Inteligencia de Grupo: Aplicación en el Balance de Carga en Redes Eléctricas de Baja Tensión», *Rev. la Esc. Perfecc. en Investig. Oper.*, n.º 40, pp. 23-42, 2016.
- [11] M. Clerc, *Particle Swarm Optimization*. 2006.
- [12] F. Glover y G.A. Kochenberger. *Handbook of Metaheuristics*. 2003.

Integración de IoT en el modelado de procesos de negocio

Luis Damian Candia¹, Anahí S. Rodríguez¹, Patricia Bazán¹, Viviana M. Ambrosi^{1,2},
Javier Díaz¹

1- LINTI – Facultad de Informática – UNLP

2- Comisión de Investigaciones Científicas de la Provincia de Buenos Aires - CIC
dcandia@linti.unlp.edu.ar, arodriguez@linti.unlp.edu.ar, pbaz@info.unlp.edu.ar,
javier.diaz@info.unlp.edu.ar, vambrosi@info.unlp.edu.ar

Resumen

En la era de Internet de las Cosas (IoT – *Internet of Things*) se calcula que para el año 2025 existirán más 26 billones de dispositivos conectados incluyendo casas, teléfonos, autos y fábricas. En este contexto es imposible imaginar que los procesos de las organizaciones no deban adaptarse a la nueva realidad, en la cual la mayoría de las actividades serán realizadas por autómatas que deberán tomar decisiones en base a la información obtenida por su red de sensores.

Así como es necesario integrar IoT con soluciones IT tradicionales donde se procesan datos transaccionales, también las soluciones orientadas a procesos de negocio (PN– *Procesos de Negocio*) y su naturaleza estática, se ven desafiadas por el mundo de IoT donde se hace presente el dinamismo de las restricciones y la heterogeneidad de las entradas.

La línea de investigación propone presentar las debilidades y fortalezas de la integración de IoT en una solución de BPM, y presentarlas en un caso de estudio. Teniendo en cuenta que la aplicación de BPM conlleva un ciclo de vida donde sus etapas integran un ciclo de mejora continua para modelado, despliegue, ejecución y monitoreo de los procesos de negocio, en este trabajo se definirá la integración de BPM e IoT para la fase de modelado, quedando como

trabajo futuro, abarcar todas las etapas del ciclo de vida de los procesos de negocio.

El caso de estudio a modelar estará basado en un proceso de negocio que incluya componentes de IoT, de la Planta Piloto Experimental de Residuos Electrónicos desarrollada en conjunto entre la UNLP, el programa E-Basura y la *International Telecommunications Union* (ITU).

Palabras clave: BP, BPM, IoT, E-Basura, Tecnologías Verdes, Desarrollo Sostenible, Green IT

Contexto

Las líneas de trabajo que se describen en el artículo se desarrollan dentro del Laboratorio de Investigación de Nuevas Tecnologías Informáticas (LINTI) [1] de la Facultad de Informática [2] enmarcadas en el Proyecto de I+D “Internet del futuro: Ciudades digitales inclusivas, innovadoras y sustentables, IoT, ciberseguridad y espacios de aprendizaje del futuro”.

Asimismo, desde el año 2009 la Facultad de Informática de la UNLP [3] y el LINTI llevan adelante un proyecto de extensión universitaria, el Proyecto E-Basura [4], que partir de fines del 2017 fue declarado Programa por la propia UNLP. Dicho programa brinda una solución interdisciplinaria a la problemática de los Residuos de Aparatos

Eléctricos y Electrónicos (RAEE) tanto para la universidad como para su entorno.

Uniendo ambas líneas de investigación se aplicará la integración a un caso de estudio real. Se incluirá la etapa diseño de un proceso de negocio de la Planta Piloto de reciclado que utilice componentes de IoT.

Introducción

Internet de las Cosas (IoT) es una tecnología emergente con un potencial enorme, que se estima va a traducirse en mayores ingresos, mayor eficiencia y menores costos para las organizaciones, pero su despliegue puede ocasionar otros inconvenientes a la hora de su integración con tecnologías preexistentes.

IoT es un fenómeno tecnológico bastante reciente. El término fue introducido por Kevin Ashton [5], investigador en el campo de la identificación por radiofrecuencia (RFID). Este autor plantea el debate sobre qué ocurriría si absolutamente todos los objetos que nos rodean nos brindaran información útil al instante. Esto cambiaría por completo nuestra forma de desenvolvemos con el mundo que nos rodea. Las computadoras podrían recopilar información sobre los objetos físicos sin intervención humana mediante el uso de tecnología de sensores y actuadores.

La gestión de los procesos de negocio es imprescindible para la administración de una organización. Weske [6] define BPM (*Business Process Management*), como la representación de las actividades de un proceso de negocio y sus restricciones de ejecución entre ellas, y que incluye conceptos, métodos y técnicas para apoyar su ciclo de vida.

El ciclo de vida de un proceso de negocios consta de las siguientes etapas: Diseño; Configuración (Definición);

Promulgación (Ejecución); Evaluación (Monitorización).

Este trabajo se centra en la etapa de diseño del proceso de negocio que se alimenta de una red de sensores. El modelado se realiza usando la notación BPMN. Esta notación es un estándar definido por OMG (*Object Management Group*) [10] y está diseñada para coordinar la secuencia de los procesos y los mensajes que fluyen entre los participantes de las diferentes actividades. BPMN [10] proporciona un conjunto de símbolos comunes para que las partes involucradas puedan comunicar los procesos de forma clara, completa y eficiente produciendo un Diagrama de Procesos de Negocio (*Business Process Diagram*, BPD).

Líneas de Investigación y Desarrollo

El futuro de IoT dependerá de cómo sus componentes puedan interactuar con los procesos de negocio o con las soluciones de IT que se ejecuten a su alrededor. Una costosa red de sensores carece de utilidad si no se construyen soluciones de software que consuman las señales. Si ese software está conformado por componentes orquestados como procesos de negocio, es natural pensar que los elementos de IoT deban ser tenidos en cuenta desde la fase de modelado de dichos procesos de negocio.

Según plantea Meyer [7], un primer desafío implica definir elementos notacionales especiales para contar con modelo orientado a IoT. Varios autores acuerdan con esto y en este trabajo se propone un enfoque posible para abordar este desafío y contar con un BPD con elementos de IoT.

En la Tabla 1 se muestran los elementos de IoT definidos por Meyer y

Sperner [8][7][11], las propuestas de diferentes autores para cada elemento y una columna con nuestra propuesta.

Caso de Estudio

Se presenta el modelo obtenido utilizando la notación propuesta para modelar un proceso de negocio de la Planta Piloto Experimental de Residuos Electrónicos UNLP - E-Basura - ITU [12]. El proceso seleccionado mide el nivel de contaminación sonora que se registra en el área de reciclado de equipamiento informático. El mismo es producido debido a la utilización de un molino de plásticos orientado al reciclado de materiales de descarte industrial y post consumo, que muele el plástico de los componentes de los RAEE.

En el diagrama del proceso para medir la contaminación sonora de la Figura 1, se observan los componentes mencionados en la Tabla 1 para representar una solución orientada a IoT en un BPD. El usuario se representa por un *Lane* llamado *Responsable del Taller*, en él se ejecutarán las tareas de recepción de alertas e informes, la entidad física se encuentra modelada como una notación de texto denominada *MedioAmbiente*. La entidad virtual es modelada en el diagrama como un repositorio de datos llamado *MedioAmbiente* desde el cual el sistema tomará información para decidir su acción, por ejemplo, en caso que la lectura supere un indicador predeterminado enviará una alerta, de acuerdo a rangos especificados previamente. El dispositivo de IoT utilizado es una placa Raspberry PI 3 Model B, la cual es modelada también como un *Lane* que consume recursos de la entidad virtual guardando la información en una base de datos y, en caso de ser necesario interactúa con el usuario. El ciclo de lectura continúa

siempre y cuando se encuentren las lecturas habilitadas en el sistema.

Formación de Recursos Humanos

BPM como modelo para la construcción de soluciones tecnológicas constituye un área de trabajo consolidada desde hace varios años y que cada vez encuentra más puntos de contacto con otros temas y líneas de investigación.

El mundo de Internet de las Cosas posee una definición extremadamente amplia. Una posible visión de este concepto es describirlo como una red de sensores interconectados, que encuentra puntos de contacto con BPM, al cual puede verse como un circuito de actividades que consumen servicios. Por otra parte, existe una divergencia en la estabilidad y unicidad de los servicios orquestados por BPM frente al entorno inestable y cambiante de una red de sensores.

Siendo IoT parte de las TI, surge la necesidad creciente de interconectar los dos mundos y hacerlos interactuar de modo que BPM aporte la estabilidad al mundo de IoT y a su vez pueda incorporar parte de su dinamismo,

El programa E-Basura de la UNLP, se encuentra ampliamente consolidado y enmarcado en la Planta Piloto con la ITU, requiere cumplir con normas internacionales que se imponen como resultado de su crecimiento. En particular, requiere adherir a la definición de KPIs (*Key Performance Indicator*, en español, Indicadores Clave de Rendimiento) indicados por la ITU [12] para el tratamiento y gestión de los indicadores clave de rendimiento medioambiental.

Por tal motivo, es válida y necesaria la conformación de un grupo de trabajo que sume esfuerzos para el crecimiento de la Planta Piloto y la aplicación de nuevas

tendencias tecnológicas con el fin de lograr madurez de los conceptos

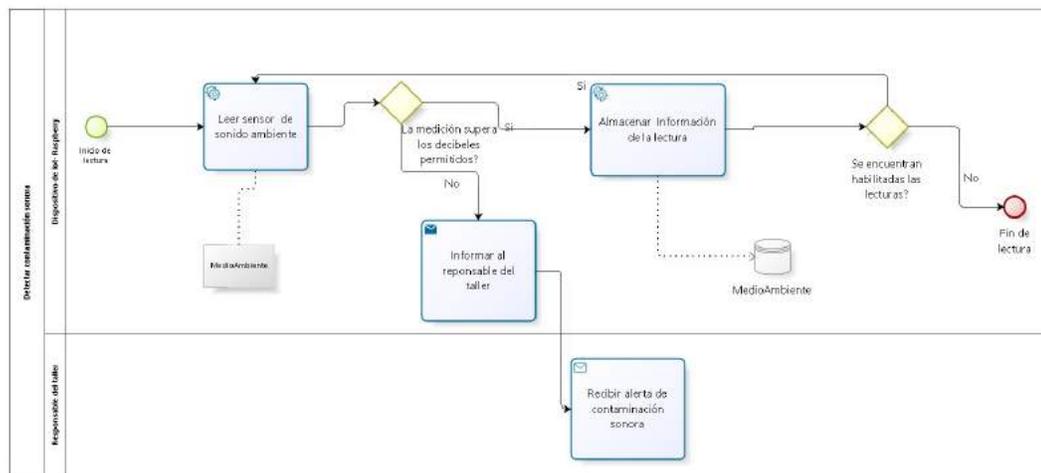


Figura 1 – BPD del proceso de lectura de contaminación sonora orientado a IoT

Resultados y Objetivos

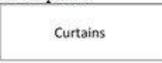
La presente línea de investigación propone integrar los conceptos y tecnologías de IoT con todo el ciclo de vida de los procesos de negocio.

En particular, se presenta una estrategia de modelado de procesos de negocio incorporando elementos de IoT. La propuesta cubre solo una fase del ciclo de vida de los procesos de negocio pero constituye un primer paso en la definición de una metodología que incorpore las características flexibles y dinámicas del ámbito de IoT a los procesos de negocio y las soluciones de IT que los despliegan.

Referencias

1. Laboratorio de Investigación de Nuevas Tecnologías Informáticas, <https://www.linti.unlp.edu.ar/>
2. Facultad de Informática, <http://info.unlp.edu.ar/>
3. Universidad Nacional de La Plata, <https://unlp.edu.ar/>
4. Proyecto E-Basura , <https://e-basura.linti.unlp.edu.ar/>
5. Ashton, K. (2009). That ‘internet of things’ thing. *RFID journal*, 22(7), 97-114.
6. Weske, M. (2012). Business process management architectures. In *Business Process Management* (pp. 333-371). Springer, Berlin, Heidelberg.
7. Meyer, S., Sperner, K., Magerkurth, C., & Pasquier, J. (2011, June). Towards modeling real-world aware business processes. In *Proceedings of the Second International Workshop on Web of Things* (p. 8). ACM.
8. Meyer, S., Ruppen, A., & Magerkurth, C. (2013, June). Internet of things-aware process modeling: integrating IoT devices as business process resources. In *International conference on advanced information systems engineering* (pp. 84-98). Springer, Berlin, Heidelberg.
9. AG, S. Internet of Things Architecture IoT-A Project Deliverable D2. 2–Concepts for Modelling IoT-Aware Processes.
10. About the Business Process Model And Notation Specification Version 2.0.2 <https://www.omg.org/spec/BPMN/About-BPMN/>
11. De Meyer, T. INTEGRATING THE INTERNET OF THINGS INTO BUSINESS PROCESS MANAGEMENT.
12. International Telecommunication Union, <https://www.itu.int>

Tabla 1 'Propuesta de Elementos de IoT en notación BPMN

Entidad IoT	Descripción	Representación BPM propuesta por distintos autores	Nuestra propuesta	Representación Grafica
Usuario	Las personas que se ven beneficiarias por la solución de IoT.	Meyer propone modelar al usuario como un proceso que envía un mensaje solicitando un recurso o un servicio. 	Se propone modelarlo como un <i>Lane</i> más que participa del proceso de lectura, e interactúa con el <i>Lane</i> del dispositivo de IoT mediante el envío de mensajes, realizando una comunicación hombre-máquina.	
Entidad Física	Elemento del mundo real que se conecta a la red de sensores y actuadores.	Meyer propone modelar la entidad física como un proceso pasivo. 	Se propone modelarlo como una notación de Texto que representa los datos involucrados en el entorno. La representación del entorno no impacta en el proceso activo.	
Entidad Virtual	Representación virtual de la entidad física.	Sperner propone modelar como un objeto de datos. 	Se propone modelarlo como un Repositorio de Datos, ya que será el encargado de almacenar los datos enviados por el dispositivo IoT.	
Dispositivo de IoT	Componente que interactúa directamente con los sensores y actuadores.	Meyer propone modelar el dispositivo de IoT, como un Lane, debido a que es un actor mas del proceso de negocio. 	En nuestro caso de estudio utilizaremos la misma notación propuesta por Meyer para modelar el dispositivo de IoT.	
Recursos y Servicios	Un recurso o servicio puede verse como una acción a realizar, como leer un dato de un sensor.	Meyer propone modelar los recursos o servicios como una tarea automática. 	En nuestro caso de estudio utilizaremos la misma notación propuesta por Meyer para modelar recursos y servicios como una actividad automática.	

Revisión con base cognitiva de un Proceso de Requisitos

Graciela D.S. Hadad^{1,2}, Jorge H. Doorn^{1,2}, María C. Elizalde¹

¹ Escuela de Informática, Universidad Nacional del Oeste

² DIIT, Universidad Nacional de la Matanza

ghadad@uno.edu.ar, jdoorn@exa.unicen.edu.ar, melizalde@uno.edu.ar

RESUMEN

En todas las etapas del desarrollo de software, están presentes variadas tareas cognitivas, inmersas en el proceso mismo. Estas tareas cognitivas han sido poco estudiadas, aún cuando afectan el desempeño de las personas intervinientes en el proyecto. Si bien todas las actividades del desarrollo de un sistema de software involucran tareas cognitivas, las correspondientes a las primeras etapas tienen una mayor presencia de las mismas. Es decir, durante el proceso de requisitos, la colaboración, la comunicación y la comprensión del contexto donde se desempeñará el sistema son partes inherentes del mismo. Es por ello que se considera que la psicología cognitiva, y particularmente la ergonomía cognitiva, pueden ayudar a lograr una mejor comprensión de las actividades humanas que se llevan a cabo en esta etapa del desarrollo de software. En el presente proyecto se planifica estudiar, revisar, evaluar y eventualmente proponer modificaciones a las heurísticas de un proceso específico de requisitos, focalizando en aspectos cognitivos de los involucrados, de manera tal de mejorar el desarrollo de las actividades del proceso y, por ende, mejorar la calidad del producto final.

Palabras clave: Ingeniería de Requisitos, Proceso de Requisitos, Ergonomía Cognitiva, Elicitación, Lenguaje Natural.

CONTEXTO

La línea de investigación que se presenta es parte de los proyectos de investigación “Tratamiento de los Factores Situacionales y la Complejidad en la Ingeniería de Requisitos” de la Universidad Nacional del Oeste (UNO) y “Reflexividad como herramienta en la

Ingeniería de Requisitos” de la Universidad Nacional de La Matanza (UNLaM).

1. INTRODUCCIÓN

En muy diversas actividades de la Ingeniería de Software, en general, y de la Ingeniería de Requisitos, en particular, es altamente recomendable utilizar heurísticas, especialmente en aquellas relacionadas con modelos construidos manualmente o en forma parcialmente automatizada. Frecuentemente estas heurísticas son adaptadas, no siempre en forma planificada, a circunstancias particulares del contexto de cada proyecto de desarrollo, tanto por razones relacionadas con el proyecto como por características de los usuarios de las mismas [1]. Sin embargo, rara vez se las cuestiona de raíz, procurando concentrar la atención en el sustento conceptual de las mismas.

Una consecuencia inmediata de lo anterior es preguntar cómo es que han sido construidas esas heurísticas y, por supuesto, la pregunta más general acerca de cómo debería ser construida cualquier heurística.

La respuesta a la segunda pregunta tiene la apariencia de ser sumamente simple y justamente esa apariencia es el origen de su dificultad intrínseca. En virtud del alto contenido de información de la mayoría de los modelos utilizados en las diferentes actividades de la Ingeniería de Requisitos, resulta ser que las heurísticas deberían guiar al ser humano en el desarrollo de las actividades cognitivas involucradas.

La mera introducción de la palabra *cognitiva* hace evidente lo erróneo de considerar simple la construcción de una heurística en ésta y en muchas otras áreas de la Ingeniería de Software.

Abordar la planificación y ordenamiento

de cualquier actividad intelectual, aún aquellas con componentes rutinarios, es difícil ya que se deben considerar las fortalezas y debilidades del ser humano en relación con dichas actividades.

El estudio del desempeño del ser humano en entornos deportivos, laborales y situaciones de todo tipo tiene larga data, aunque en la mayoría de los casos orientados a problemáticas de actividades motrices y uso de herramientas. Estos estudios han dado lugar a la disciplina denominada *ergonomía* [2, 3], también conocida como *factores humanos*.

Posteriormente, los estudios relacionados con los factores humanos han ido adquiriendo un perfil cada vez más relacionado con aspectos sensoriales y cognitivos [4], habiéndose centrado principalmente en las interacciones con sistemas complejos [5]. La introducción de estos enfoques en los sistemas de procesamiento de datos se produjo en el área de visualización y diseño de interfaces [4, 6]. Aún hoy en día, la mayoría de estos estudios siguen estando en las cercanías de aspectos sensoriales y no verdaderamente cognitivos.

Todo lo anterior hace evidente que este es un terreno prácticamente inexplorado cuyas características son simultáneamente muy promisorias y desafiantes. Las promesas residen en el hecho que al poner en el centro del problema las peculiaridades de la cognición humana se avizoran importantes mejoras en la forma de abordar la elicitación de conocimiento y en su posterior elaboración y registro, mientras que los desafíos residen en que los aportes de las ciencias cognitivas tienen aún pocos resultados de importancia pragmática.

Como consecuencia, el presente proyecto de investigación tiene esencialmente características exploratorias [7, 8]. En otras palabras, se trata de elaborar hipótesis que permitan guiar futuros proyectos en los que las mismas sean confirmadas, mejoradas o refutadas. Usualmente se suele considerar que la definición de las heurísticas es una tarea sencilla y casi irrelevante. Definitivamente este no es el caso. En diversos proyectos

previos, se ha comprobado lo irreductible que son algunos de las debilidades conocidas en el proceso de requisitos, especialmente aquellas vinculadas con los problemas de completitud [9, 10, 11, 12].

Las cuatro fuentes básicas de hipótesis [7] son: i) marcos conceptuales o teóricos, ii) referencias bibliográficas, iii) investigaciones previas, y iv) intuiciones, sospechas y experiencias de la vida diaria. Se estima que en el presente proyecto se deberá recurrir a todas ellas.

La necesidad de tener en cuenta las limitaciones del ser humano en cuanto a su capacidad de comprensión y análisis de problemas de toda naturaleza ha estado presente en la informática desde etapas muy tempranas y sigue estando. Posiblemente el ejemplo más relevante en este sentido sea la fuerte recomendación de guiarse por enfoques top-down en diversas áreas y en utilizar el método de los refinamientos sucesivos en particular [13]. Pese a lo valioso de estas contribuciones es relevante destacar que en el artículo original de Niklaus Wirth no hay mención alguna a la necesidad de concentrar la atención del programador en unos pocos aspectos en cada momento del desarrollo. Tampoco se indica que se esté enfrentando la creación de un programa para atender una necesidad perfectamente comprendida y especificada.

Al considerar el traslado de estos enfoques al área de la Ingeniería de Requisitos, en su muy valiosa contribución [14], Michael Jackson enfáticamente aconseja no usar enfoques top-down en virtud que hacen necesario fragmentar el sistema a estudiar en el peor momento posible, que es cuando menos se lo conoce. Es decir, en el momento en que existe mayor probabilidad de cometer errores por la falta de comprensión acabada del problema. Nuevamente se encuentra aquí un sustento, algo más explícito, de atender las necesidades cognitivas de las personas intervinientes.

Otros ejemplos similares son el uso de abstracción mediante lenguajes de modelado del dominio [15] y la descomposición mediante diagramas de flujos de datos [16].

Se pueden agregar otros muchos y muy variados ejemplos, en los que se pueden observar innovaciones relacionadas con aspectos cognitivos. Sin embargo, en casi todos los casos, las cuestiones cognitivas tienen un rol de telón de fondo. Es decir, son poco analizadas y su influencia sobre las innovaciones introducidas no ha sido efectivamente probada.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Una somera revisión de la literatura en lo que se refiere a las heurísticas que soportan fuertemente la gran mayoría de los métodos de la Ingeniería de Software muestra que las mismas están básicamente orientadas a describir el contenido de los modelos que a ayudar a elaborar conceptualmente ese contenido.

Si se analiza el Proceso Unificado en su flujo de trabajo Requisitos, éste propone la captura de requisitos centrándose en el modelado del dominio, del negocio y de los casos de uso, con heurísticas que atienden puramente a construirlos [17]. Por ejemplo, el modelo de dominio se describe mediante diagramas de clase UML cumpliendo con: “las clases del dominio aparecen como i) objetos del negocio que representan cosas que se manipulan en el negocio, ii) objetos del mundo real y conceptos que el sistema usará, y iii) sucesos que ocurrirán o han ocurrido en el entorno de trabajo” [17, pág.113]. Para la descripción de casos de uso, el Proceso Unificado menciona: “i) Debe definir el estado inicial como precondition, ii) Debe definir los posibles estados finales, iii) La interacción del sistema con los actores y qué cambios producen”, entre otros puntos [17, pág. 149]. En general, se observa que el Proceso Unificado no propone ninguna ayuda acerca de cómo seleccionar elementos del dominio, del negocio o del sistema, tales como clases, actores e interacciones relevantes, o como descartar lo superfluo.

En general, la literatura referida al Desarrollo Dirigido por Modelos (MDD) [18] considera que los dos primeros modelos CIM (Computational Independent Model) y PIM

(Platform Independent Model) en la cadena de transformación no pueden ser obtenidos en forma automatizada a partir de reglas de construcción (caso de CIM) ni de reglas de derivación del modelo previo (caso de PIM), sino que en ambos casos deben obtenerse a través de heurísticas, las que atienden aspectos netamente relacionados con el meta-modelo que los soporta sin guiar en cómo obtener el contenido. Es decir, en este paradigma aún no se han podido establecer reglas que automaticen la construcción de estos modelos ni que colaboren con la determinación de qué tipo de información del contexto de aplicación es necesario capturar, siendo considerada hoy en día una actividad “artesanal” [15]. Como declara Booch et al. [19] en su manifiesto de MDD solo se automatiza la transformación de modelos que no dependan de la “ingenuidad humana”.

Numerosos autores reportan que incrementa la probabilidad de éxito de un proyecto de software el estudio del dominio del problema y luego que los requisitos de ese sistema se plasmen en modelos creados en Lenguaje Natural [20, 21]. Las mejoras registradas en la calidad de los requisitos tienen su fundamento en que estas representaciones, tales como glosarios, casos de uso y escenarios, promueven la comunicación entre todos los involucrados en un proyecto de desarrollo de software y facilitan la validación de los requisitos elaborados. Una revisión hecha por Rolland et al. [22] muestra que, de doce enfoques propuestos en la literatura en el ámbito de la Ingeniería de Requisitos, todos ellos usan una notación de texto para describir escenarios, que en algunos casos se combinan con otros medios, tales como gráficos o imágenes. En un estudio posterior sobre la práctica en Ingeniería de Requisitos [23], se concluyó que el 51% de las organizaciones (sobre un total de 194) usa representaciones informales (por ejemplo, el lenguaje natural) y el 27% semi-formales, quedando los modelos formales relegados a un uso de apenas el 7% de las organizaciones.

En proyectos de investigación previos, se ha desarrollado una estrategia de IR basada en

modelos en lenguaje natural [24]. Aún cuando esta estrategia puede considerarse suficientemente madura, es deseable que la misma pueda ser mejorada en varios aspectos. Este proyecto está enmarcado en esa estrategia.

La estrategia procura alcanzar una profunda comprensión del dominio de la aplicación para definir una solución óptima a través de un sistema de software. Es decir, la estrategia presenta dos grandes etapas bien distinguibles: una de aprendizaje y la otra de definición. Cuando hay un conocimiento previo del dominio de la aplicación la primera fase se convierte en un proceso confirmatorio.

Los modelos que se utilizan en esta estrategia [24] son:

- un modelo léxico, LEL (Léxico Extendido del Lenguaje), el cual describe el vocabulario utilizado en el dominio de la aplicación,
- un modelo organizacional que describe los procesos actuales del negocio utilizando, denominado Escenarios Actuales,
- un modelo organizacional que describe los procesos del negocio proyectados en base al sistema de software a desarrollar, denominado Escenarios Futuros.

La mayoría de las heurísticas de esta estrategia de IR adolecen de las mismas debilidades que se han mencionado para el Proceso Unificado y para MDD, en el sentido que están fuertemente orientadas al contenido de los modelos que se construyen, con pocas pautas sobre cómo capturar y analizar la información a ser registrada en ellos.

3. RESULTADOS OBTENIDOS/ESPERADOS

En experiencias preliminares [25, 26] se han aplicado nociones de ergonomía cognitiva a la construcción del modelo Léxico Extendido del Lenguaje, poniendo especial énfasis en el uso de una heurística radicalmente diferente de la usual y orientada a facilitar la percepción del significado de los términos por parte del ingeniero de requisitos. Los resultados de este estudio fueron notoriamente prometedores, ya que se logró

una cantidad de términos notoriamente superior a la que se había obtenido anteriormente.

En la Tabla 1 se resumen los datos recogidos en una de estas experiencias [27]. Las omisiones indicadas en la misma están calculadas comparando el LEL obtenido con una versión anterior del mismo LEL creado por otros autores, habiendo sido ambas experiencias controladas en el sentido que ambos trabajos fueron realizados con el mismo material y con absoluto desconocimiento del otro trabajo y de la intención de hacer una comparación posterior.

Tabla 1. Omisiones en la versión de un LEL respecto a otra versión creada aplicando nociones cognitivas

Tipo de Término	Omisiones
Sujetos	17 %
Objetos	27 %
Verbos	62 %
Estados	44 %
Total	39 %

Es notable el déficit en la detección de los verbos con las heurísticas preexistentes. También se destaca que siempre se tuvo la suposición, poco fundada, que la mayor debilidad de la actividad de construcción del LEL residía en la identificación apropiada de los estados, hecho que parece no ser cierto.

Estos resultados son meramente indicativos, por dos razones: i) no tienen significación estadística, y ii) no se han establecido hipótesis que vinculen las visiones cognitivas con la heurística utilizada.

Tal como se mencionó en la introducción, desde el punto de vista epistemológico, el presente proyecto se puede catalogar como una investigación esencialmente exploratoria, en la que se aspira a: i) detectar las principales dificultades cognitivas que enfrenta el ingeniero de requisitos al elicitar y modelar el conocimiento que se adquiere del universo del discurso y al concebir el conjunto de servicios que deberá prestar el sistema de software a ser desarrollado, y ii) proponer nuevas heurísticas que no sólo describan los modelos a ser confeccionados sino que efectivamente contribuyan a facilitar la elaboración de la información que se registre en los mismos.

Obviamente, se espera que estos resultados puedan luego ser contrastados en investigaciones confirmatorias.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto de la UNLaM participan tres investigadores, dos de ellos en formación. En el proyecto de la UNO participan en este tema otros dos investigadores, uno de ellos en formación.

La línea de investigación presentada aquí es parte directa de las tesis de maestría de la Ing. Renata Guatelli y de las tesis de doctorado de la Mg. Gladys Kaplan y la Ing. Andrea Vera, y colabora con la tesis de doctorado de la Lic. María Pepe.

5. BIBLIOGRAFÍA

- [1] Hadad, G.D.S., Doorn, J.H., Ledesma, V.A. (2018) Dynamic Situational Adaptation of a Requirements Engineering Process, en: Encyclopedia of Information Science and Technology, Fourth Edition, IGI Global, cap. 646, pp. 7422-7434.
- [2] Murrell, K. (1965) Ergonomics Man in His Working Environment, Springer, pp 103-124.
- [3] Yong Zhou, Y., (1974) Applied Ergonomics Handbook, Elsevier, pp. 1-7.
- [4] van der Veer, G. (2008) Cognitive Ergonomics in Interface Design - Discussion of a Moving Science, Journal of Universal Computer Science, Vol. 1, N° 16, pp. 2614-2629.
- [5] Gersh, J.R., McKneely, J.A., Remington, R.W. (2005) Cognitive engineering: Understanding human interaction with complex systems, Johns Hopkins APL Technical Digest, Vol. 26, N° 4, pp. 377-382.
- [6] Kylie, M., Gomes, K.M., Riggs, S.L. (2017) Analyzing Visual Search Techniques using Eye Tracking for a Computerized Provider Order Entry (CPOE) Task, Human Factors and Ergonomics Society Annual Meeting, pp. 691-695.
- [7] Hernández Sampieri, R., Fernández Collado, C., Baptista Lucio, M.P. (2010) Metodología de la Investigación, 5ta. ed., Mc Graw Hill.
- [8] Popper, K. (1963) Conjectures and Refutations: The Growth of Scientific Knowledge, Routledge.
- [9] Doorn, J.H., Ridao, M.N. (2003) Completitud de Glosarios: Un estudio experimental, VI Workshop on Requirements Engineering, pp. 317-328.
- [10] Ridao, M.N., Doorn, J.H. (2006) Estimación de Completitud en Modelos de Requisitos Basados en Lenguaje Natural, XI Workshop on Requirements Engineering, pp. 146-157.
- [11] Doorn, J.H., Ridao, M.N. (2008) Completeness Concerns in Requirement Engineering, en: Encyclopedia of Information Science and Technology, Second Edition, IGI Global.
- [12] Hadad, G.D., Litvak, C. S., Doorn, J.H., Ridao, M. (2015) Dealing with Completeness in Requirements Engineering, en: Encyclopedia of Information Science and Technology, Third Edition, IGI Global.
- [13] Wirth, N. (1971) Program Development by Stepwise Refinement. Communications of the ACM, Vol. 14, N° 4, pp. 221-227.
- [14] Jackson, M., A. (1995) Software Requirements & Specifications: a lexicon of practice, principles and prejudices, Addison Wesley.
- [15] Pons, C., Giardini, R., Pérez, G., (2010) Desarrollo de Software Dirigido por Modelos: Conceptos teóricos y su aplicación práctica, Editorial de la Universidad Nacional de La Plata / McGraw-Hill Educación.
- [16] Jourdon, E. (1989) Modern Structured Analysis, Englewood Cliffs, NJ, Yourdon Press/Prentice Hall.
- [17] Jacobson, I., Booch, G., Rumbaugh, J., (2000) El Proceso Unificado de Desarrollo de Software, Pearson Educación, Madrid.
- [18] Karow, M., Gehlert, A. (2006) On the Transition from Computation Independent to Platform Independent Models, 12th American Conference on Information Systems, Mexico.
- [19] Booch, G., Brown, A., Iyengar, S., Rumbaugh, J., Selic, B. (2004) An MDA Manifesto, en Frankel, D. and Parodi J. (eds) The MDA Journal: Model Driven Architecture Straight from the Masters.
- [20] Ryan, K. (1993) The Role of Natural Language in Requirements Engineering, IEEE International Symposium on Requirements Engineering, San Diego, pp. 240-242.
- [21] Loucopoulos, P., Karakostas, V. (1995) System Requirements Engineering, McGraw-Hill, Londres.
- [22] Rolland, C., Souveyet, C, Ben Achour, C. (1989) Guiding Goal Modeling Using Scenarios, IEEE TSE, Vol. 24, N° 12, pp. 1055-1071.
- [23] Neill, C.J., Laplante, P.A. (2003) Requirements Engineering: The State of the Practice, IEEE Software, Vol.20, N° 6, pp.40-45.
- [24] Leite, J., Doorn, J., Kaplan, G., Hadad, G., Ridao, M. (2004) Defining System Context using Scenarios. Perspectives on Software Requirements, Kluwer Academic Publishers, cap.8, pp.169-199.
- [25] Litvak, C.S., Hadad, G.D.S., Doorn J.H. (2017) Nominalizations in Requirements Engineering Natural Language Models, en: Encyclopedia of Information Science and Technology, Fourth Edition, IGI Global.
- [26] Litvak, C.S., Hadad, G.D.S., Doorn, J.H. (2016) Procesamiento de Lenguaje Natural para Estudiar Completitud de Requisitos. Workshop de Investigadores en Ciencias de la Computación, pp. 498-502.
- [27] Litvak, C.S., Doorn J.H. (2016) Aporte del Procesamiento de Lenguaje Natural a la Ingeniería de Requisitos, 4to Congreso Nacional de Ingeniería Informática / Sistemas de Información, pp.463-467.

Estrategias para la clasificación de contenido y usuarios de Foros de Discusión Técnicos

Gabriela Aranda, Nadina Martínez Carod, Valeria Zoratto, Alejandra Cechich,
Facundo Otermin Sánchez, Carina Noda, Mauro Sagripanti

Grupo de Investigación en Ingeniería de Software del Comahue (GIISCO)
<http://giisco.uncoma.edu.ar>

Facultad de Informática. Universidad Nacional del Comahue
Buenos Aires 1400, (8300) Neuquén

Contacto: {gabriela.aranda, nadina.martinez, valeria.zoratto, alejandra.cechich}@fi.uncoma.edu.ar

RESUMEN

Los foros de discusión son utilizados por muchos técnicos informáticos para plantear dudas y pedir sugerencias para resolver algún problema particular. Para ello, formulan una pregunta a partir de la cual se abre un hilo de discusión (thread), en el que suelen participar varios usuarios que analizan el escenario y proponen una o más soluciones al problema en cuestión. De esta manera, los foros se han convertido en plataformas colaborativas donde el conocimiento se explicita a la vez que se comparte.

Dado que existen muchos foros sobre las mismas temáticas (lenguajes de programación, aplicaciones específicas, etc.), es posible encontrar en la Web muchos hilos de discusión en diferentes foros que están relacionados al mismo problema. Cuando un técnico informático tiene un problema específico, suele utilizar un motor de búsqueda multi-propósito que le devuelve una lista extensa de páginas de varios tipos (blogs, foros, artículos, etc.), luego el técnico necesita navegar por varias páginas hasta descubrir cuál es la que describe un problema más parecido al que tiene, y encontrar (si existe) una solución que pueda satisfacerle.

Para facilitar esta tarea periódica de los técnicos informáticos, nuestro proyecto tiene como objetivo la implementación de una herramienta que recupere la información disponible en hilos de discusión de foros técnicos de manera automática, y que a partir de un análisis basado en un modelo de calidad pertinente, permita clasificar dicha información y entregar a los usuarios un ranking de posibles soluciones para problemas recurrentes.

Palabras Clave

Foros de discusión, Reuso de conocimiento, Modelos de Calidad.

CONTEXTO

Nuestra línea de investigación se denomina “Reuso de Conocimientos en Foros de Discusión II” y forma parte del programa de investigación “Desarrollo de Software Basado en Reuso – Parte II”, de la Universidad Nacional del Comahue, con período de vigencia 2017-2020. El programa mencionado extiende el programa “Desarrollo de Software Basado en Reuso” realizado durante el período 2013-2016.

1. INTRODUCCION

La disciplina de Recuperación de Información (Information Retrieval) surge en la década de 1950 [12], ante la necesidad de procesar y reutilizar la información almacenada en grandes volúmenes. A partir de ese momento, este campo de investigación ha madurado y han surgido importantes aportes. Por un lado, varios proyectos se han enfocado en utilizar la información recuperada de documentos específicos, mientras que otros han desarrollado técnicas para generación automática de tesauros (lista de sinónimos, en conjunto con lista de antónimos, etc.) para su uso en distintos tipos de consultas. En general, la recuperación de información se realiza a partir de la consulta de un usuario. Luego, las posibles respuestas se organizan de acuerdo a un ranking que evalúa el grado de relevancia de cada respuesta con dicha consulta.

Si bien el conocimiento en la Web se encuentra diseminado en distintos tipos de sitios y documentos, nuestro proyecto pone el foco en los foros de discusión, que se caracterizan por ser herramientas colaborativas con grandes volúmenes de información, accesibles a la comunidad en general como fuente de consulta. En dichos foros se intercambia conocimiento entre los miembros de una comunidad de usuarios que comparte intereses y características similares.

En general, la mayoría de los métodos automáticos de IR se basan en analizar la ocurrencia de palabras en colecciones de documentos, a partir de lo cual se construyen listas de palabras fuertemente relacionadas. El principal problema detectado en estas técnicas es que no todas las palabras relacionadas con una palabra de consulta son significativas en el contexto de la consulta. Este es un

aspecto fundamental considerado en nuestro proyecto.

Dado que en la Web existen muchos foros de discusión sobre la misma temática, es posible hallar preguntas y respuestas similares diseminadas en varios de ellos, por lo que generalmente es necesario navegar por varios hilos hasta dar con una solución adecuada. Incluso, a veces es necesario considerar otras características de calidad para evaluar distintas soluciones [1][3][8].

Existen varias propuestas de reuso de conocimiento disponible en foros de discusión: Por ejemplo Chen y Persen [2] implementan un sistema recomendador que busca y agrupa mensajes con contenido similar. Por otro lado, Helic y otros [4], propone clasificar los mensajes de foros de acuerdo a una jerarquía de temas preestablecida. Luego, el enfoque de Nicoletti [17] clasifica los mensajes acorde a una jerarquía de temas obtenido de Wikipedia. Finalmente, existen propuestas de generación de algoritmos de ranking basados en la calidad de los atributos, como el que se plantea en [11].

En base a estos antecedentes, nuestro proyecto tiene como objetivo principal favorecer el reuso de la información contenida en conversaciones existentes en foros de discusión de la Web, con el valor agregado de un análisis de calidad de las fuentes de información. Además, se ha experimentado tanto con la aplicación de algoritmos de análisis de lenguaje natural como de aprendizaje automático, y se está evaluando la aplicación de *sentiment analysis* para mejorar las búsquedas. Por ejemplo, el análisis del lenguaje natural permite analizar el tipo de fragmento dentro de un hilo de discusión [10]. Teniendo esto en cuenta, nuestro proyecto está enfocado en determinar un ranking de soluciones posibles, y cada línea de

investigación dentro del proyecto lo hace desde ópticas diferentes.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Como se ha mencionado, este proyecto de investigación, denominado “Reuso de Conocimientos en Foros de Discusión – Parte II”, está enmarcado en el Programa de Investigación “Desarrollo de Software Basado en Reuso – Parte II”, de la Universidad Nacional del Comahue, con período de vigencia 2017-2020.

Dicho programa extiende la tarea realizada entre 2013 y 2016 en el Programa “Desarrollo de Software Basado en Reuso”. Respecto a este proyecto en particular, el objetivo es extender los estudios realizados sobre reuso de conocimiento en foros de discusión técnicos, incorporando la definición de métodos y algoritmos de recomendación para la asistencia inteligente a usuarios en la búsqueda de soluciones a preguntas recurrentes. Por otra parte, el programa está conformado por otros dos subproyectos que profundizan en las temáticas de Reuso Orientado al Dominio y Reuso Orientado a Servicios.

El programa “Desarrollo de Software Basado en Reuso – Parte II” está desarrollado por el Grupo de Ingeniería de Software de la Universidad Nacional del Comahue, (GIISCo), formado por docentes y estudiantes de la Facultad de Informática de la Universidad Nacional del Comahue, y cuenta con la asesoría y colaboración de otras universidades. En particular, este proyecto se lleva a cabo con la colaboración de la Facultad de Ciencias Exactas de la Universidad Nacional del Centro de la Provincia de Buenos Aires. Aunque el objetivo del Grupo GIISCo es brindar soporte en investigación y transferencia de tópicos

relacionados con la Ingeniería de Software, el proyecto también involucra a docentes pertenecientes a otras áreas de la Facultad, como Programación y Teoría de la Computación, lo que permite abordar la investigación desde ópticas diferentes, enriqueciendo el desarrollo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Como antecedentes de este proyecto de investigación, en el año 2013 se presentó un modelo de calidad para foros de discusión en base a modelos de calidad de datos e información en la Web y estándares para la calidad de datos software [9]. La validación de los atributos y subatributos de dicho modelo se realizó mediante encuestas [13]. Durante 2014 se implementó una herramienta para la recuperación de información de foros de discusión técnicos y su análisis mediante un conjunto preliminar de métricas de calidad, a partir del cual se propone un ranking de soluciones posibles para una pregunta. Dicha herramienta fue aplicada en varios casos de estudio con hilos de discusión reales y algunos de sus resultados están presentados en [27].

Entre 2015 y 2016 se avanzó en el análisis de casos de estudio a partir de una cadena de búsqueda y en el estudio del orden esperado comparado con el orden obtenido por medio de las herramientas de análisis de texto [15][16]. Para ello se utilizó la herramienta Lucene, con mecanismos personalizados para establecer *stopwords* (palabras no significativas de búsqueda) propias del dominio. En 2017, se aplicaron estas técnicas en combinación con la base de datos léxica WordNet [24], cuyos resultados preliminares fueron presentados en [28].

Esta línea de investigación se sigue desarrollando en una tesis de doctorado en la cual se evalúa distintas funciones de las bases de datos léxicas [25] para la búsqueda de mensajes relacionados a una pregunta particular.

Por otra lado, se continúan evaluando técnicas de Data Mining y modelos de aprendizaje automático supervisados y no supervisados [18][19], así como técnicas y herramientas de PLN [21] que puedan ser combinadas con las ya aplicadas.

Otra línea en marcha se enfoca en el rol de los usuarios activos de un foro (los que participan compartiendo opiniones y experiencias). Dicho análisis tiene el objetivo de incluir nuevas métricas de calidad en el recomendador de hilos de discusión en construcción. Bajo esta premisa, se han estudiado las propuestas [20] [23] [22] y se está trabajando en una tesina, a partir de una estrategia empírica basada en la observación de hilos de discusión reales obtenidos de la web.

4. FORMACION DE RECURSOS HUMANOS

El proyecto avanza en la línea del proyecto comenzado en 2013, cuyo objetivo era definir un modelo de calidad a partir de información contenida en foros de discusión técnicos.

Actualmente, el proyecto se encuentra conformado por un grupo de docentes, asesores y alumnos de las áreas de Ingeniería en Sistemas, Programación y Teoría de la Computación, trabajando en forma colaborativa e interdisciplinaria.

Las personas que colaboran, asesoran y forman parte del proyecto son:

- Dos docentes investigadores del Departamento de Programación, con dedicación exclusiva, ambos con título de Doctor en Informática.

- Un docente investigador del Departamento de Programación, con una beca doctoral otorgada por el CONICET.
- Dos docentes investigadores con dedicación simple, de los Departamentos de Ingeniería de Computadoras y de Programación.
- Tres estudiantes de Licenciatura en Ciencias de la Computación que están desarrollando sus tesis de grado dentro del proyecto.
- Una docente del Departamento de Teoría de la Computación de la misma Facultad, que está desarrollando su tesis de doctorado sobre técnicas de análisis de lenguaje natural, asesorando en temas de aprendizaje automático y lenguaje natural.
- Una docente investigadora externa, perteneciente al Instituto Superior de Ingeniería del Software (ISISTAN) de la Universidad Nacional del Centro de la Provincia de Buenos Aires. Dicha docente tiene un doctorado y experiencia en modelado de usuarios, sistemas de recomendación y Recuperación de Información.

La conformación del equipo con docentes de distintos departamentos, sumado a la asesoría externa mencionada, permite trabajo cooperativo dentro de un grupo interdisciplinario. Además, la incorporación de estudiantes de la Facultad amplía los posibles tipos de desarrollo relacionados a la temática del proyecto.

5. BIBLIOGRAFIA

- [1] ISO/IEC 25012:2008, Software product Quality Requirements and Evaluation (SQuaRE): Data quality model. 2008.

- [2] W. Chen, R. Persen (2009), “A Recommender System for Collaborative Knowledge”.
- [3] C. Calero, A. Caro, M. Piattini (2008), “An Applicable Data Quality Model for Web Portal Data Consumers”, *World Wide Web*, vol. 11, no. 4, pp. 465-484.
- [4] D. Helic, N. Scerbakov (2003), “Reusing Discussion Forums as Learning Resources in WBT Systems”.
- [5] I. Rafique et al(2012), “Information Quality Evaluation Framework: Extending ISO 25012 Data Quality Model”, *International Journal of Computer and Information Sciences*, vol.6.
- [6] R. Wang, D. M. Strong (1996), “Beyond accuracy: What data quality means to data consumers”, *Journal of Management Information Systems*, vol. 12, no. 4, pp. 5-33.
- [7] Smith y Duffy (2001), Re-using knowledge: why, what and where. En *Proceedings de 2001 International Conference on Engineering Design*, Glasgow.
- [8] P. Di Maio (2009), Toward Pragmatic Dimensions of Knowledge Reuse and Learning on the Web. *Proceedings of I-KNOW'09 and I-SEMANTICS'09*, Graz, Austria.
- [9] G. Aranda, N. Martínez Carod, P. Faraci, A. Cechich. *Hacia un framework de evaluación de calidad de información en foros de discusión técnicos*. ASSE 2013,
- [10] A. Tigelaar, R. Op Den Akker and D. Hiemstra, *Automatic summarisation of discussion fora*, *Natural Language Engineering*, ISSN 1469-8110, Vol 16, Issue 02, pp. 161-192, 2010.
- [11] H. Kuna, et al. , *Generación de un Algoritmo de Ranking para Documentos Científicos del Área de las Ciencias de la Computación*, , CACIC 2013, XIX pp. 787-796, 2013.
- [12] Singhal,. *Modern information retrieval: A brief overview*.IEEE Data Eng. Bull., 2001, vol. 24, no 4, p. 35-43
- [13] N.Martínez Carod et al. *Análisis de la información presente en foros de discusión técnicos*. In CACIC 2013, pp. 847- 856, 2013.
- [14] G. Aranda, N. Martínez-Carod, S. Roger, P. Faraci, and A. Cechich. *Una herramienta para el análisis de hilos de discusión técnicos*. In CACIC 2014, pages 803 - 812, Oct. 2014.
- [15] V. Zoratto, G. Aranda, S. Roger, A. Cechich, *Análisis de estrategias para clasificar contenidos en foros de discusión: Un caso de estudio*, ASSE 2015, pp. 176-190.
- [16] V. Zoratto, G. Aranda, S. Roger, A. Cechich, *Analyzing Discussion Forums ThreadsAbout Java Programming Language Usage*, *Electronic Journal of SADIO*, 2016 .ISSN (versión online): 1514-6774. En revisión. Publicación estimada Noviembre 2016.
- [17] M. Nicoletti, S. Schiafino, and D. Godoy. *Mining interests for user profiling in electronic conversations*. *Expert Syst. Appl.* , 40(2):638-645, Feb. 2013.
- [18] I. Witten, E. Frank and M. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier. 2011
- [19] Bing Liu. *Web Data Mining. Exploring Hyperlinks, Contents, and Usage Data*. Springer. 2008
- [20] M. Lui and T. Baldwin. *Classifying user forum participants: Separating the gurus from the hacks, and other tales of the internet*. In *Proceedings of Australasian Language Technology Association Workshop* , pages 49-57, 2010.
- [21] C. D. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, MA, USA, 1999.
- [22] T. Hecking, I. Chounta, and H. U. Hoppe. *Investigating social and semantic user roles in MOOC discussion forums*. In *LAK*, pages 198-207. ACM, 2016.
- [23] S. Bhatia and P. Mitra. *Classifying user messages for managing web forum data*. In Z. G. Ives and Y. Velegrakis, editors, *WebDB* , pages 13-18, 2012
- [24] G. A. Miller, R. Beckwith, C. D. Fellbaum, D. Gross, K. Miller. 1990. *WordNet: An online lexical database*. *Int. J. Lexicograph.* 3, 4, pp. 235–244.
- [25] A. Gangemi, R. Navigli, P. Velardi. *The OntoWordNet Project: Extension and Axiomatization of Conceptual Relations in WordNet*, In *Proc. of ODBASE 2003*, Catania, Sicily (Italy), 2003, pp. 820–838.
- [26] R. Navigli, S. P. Ponzetto. *BabelNet: Building a Very Large Multilingual Semantic Network*. *Proc. of the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)*, Uppsala, Sweden, July 11–16, 2010, pp. 216–225.
- [27] N. Martínez Carod, P. Faraci, G. Aranda *Análisis de métricas de calidad en foros de discusión técnicos*, CACIC 2017, pp.650-659
- [28] V. Zoratto, N. Martínez Carod, F. Otermín, G. Aranda: *Análisis de estrategias para clasificar contenidos en foros de discusión*, CACIC 2017, pp. 640-649

Generación automática de API REST a partir de API Java, basada en transformación de Modelos (MDD).

Ariel Arsaute, Fabio Zorzan, Marcela Daniele, Ariel González, Mariana Frutos

Departamento de Computación, Facultad de Ciencias Exactas, Físico-Químicas y Naturales,
Universidad Nacional de Río Cuarto
Ruta 36 Km. 601 –CP 5800 - Río Cuarto – Córdoba - Argentina Tel. (0358) 4676235
{arsaute, fzorzan, marcela, agonzalez, mfrutos}@dc.exa.unrc.edu.ar

RESUMEN

En la actualidad existe una gran cantidad de proyectos o aplicaciones que disponen de una API REST, una nueva opción o estilo de uso de los Servicios Web (Web Services, WS), para la creación de servicios profesionales. Twitter, YouTube, los sistemas de identificación con Facebook, y cientos de empresas generan negocio gracias a REST y las APIs REST. REST es una interfaz entre sistemas que usa HTTP para obtener datos o generar operaciones sobre esos datos en todos los formatos posibles, como XML, JSON, HTTP, etc. En los últimos años logró un gran impacto en la web que prácticamente logró desplazar a SOAP y las interfaces basadas en WSDL por tener un estilo bastante más simple de utilizar y sobre todo por su eficiencia.

La línea de investigación propone un mecanismo de generación de API REST a partir de versiones existentes de API Java, en el contexto del desarrollo dirigido por modelos (Model-Driven Development, MDD), para la construcción de WSs. Aplicar esta técnica mediante la transformación de modelos se diferencia de otras formas convencionales, las cuales se basan en generar un AST (Abstract Syntax Tree) mediante algún parser de JAVA. Además, nuestra propuesta permitirá generar código hacia distintas implementaciones de WS REST a partir de un modelo JAVA.

Palabras clave: MDD, QVT Operacional, API REST

CONTEXTO

La línea de investigación presentada en este trabajo se desarrolla en el marco del proyecto “Ingeniería de Software. La Transformación de Modelos aplicada a la Mejora continua de Procesos de Desarrollo de Software”, perteneciente a los Proyectos y Programas de Investigación (PPI) de la Secretaría de Ciencia y Técnica de la Universidad Nacional de Río Cuarto. Además esta línea de investigación se enmarca dentro de un trabajo de tesis de Maestría en Ingeniería de Software.

1. INTRODUCCIÓN

El desarrollo dirigido por modelos (MDD) [1] es una metodología de desarrollo de software que se centra en la creación y explotación de modelos de dominio. Es decir, representaciones abstractas de los conocimientos y las actividades que rigen un dominio de aplicación particular. MDD tiene como objetivo aumentar la productividad mediante la maximización de la compatibilidad entre los sistemas, a través de la reutilización de modelos estandarizados, simplificando el proceso de diseño (a través de modelos de patrones de diseño que se repiten en el dominio de aplicación), y promoviendo la comunicación entre los individuos y equipos que trabajan en un sistema por medio de una estandarización de la terminología y las mejores prácticas utilizadas en el dominio de aplicación. Los modelos son desarrollados contemplando comunicaciones entre los gerentes de producto, diseñadores, miembros del equipo

de desarrollo y usuarios del dominio de la aplicación.

El Object Management Group (OMG) [2] tiene como iniciativa de MDD la arquitectura dirigida por modelos (Model Driven Architecture, MDA) [3], que es un acercamiento al diseño de software. Bajo la metodología MDA la funcionalidad de un sistema es definida en primer lugar como un modelo independiente de la plataforma (Platform-Independent Model, PIM). Dado un modelo de definición de la plataforma (Platform Definition Model, PDM), el modelo PIM puede traducirse entonces a uno o más modelos específicos de la plataforma (Platform-Specific Models, PSM) para la implementación correspondiente, usando diferentes lenguajes específicos del dominio, o lenguajes de propósito general como Java, C#, Python, entre otros. La traducción entre un modelo PIM y modelos PSM se realiza normalmente utilizando herramientas automatizadas, como lo son las herramientas de transformación de modelos, por ejemplo, aquellas que cumplen con el estándar OMG denominado Query/View/Transformation, (QVT) [4]. Además, se propone las automatizaciones de las transformaciones entre modelos y de la generación de código, centrando el proceso de desarrollo de software en las tareas esencialmente de modelado

MDA está relacionado con múltiples normas, incluyendo el lenguaje Unificado de Modelado (Unified Modeling Language, UML) [5], que se ha convertido en un estándar. UML sigue el paradigma de orientación a objetos y permite la descripción de aspectos tanto estáticos como dinámicos de sistemas de software. Más que un lenguaje es un conjunto de lenguajes, en su mayoría notaciones gráficas, soportados por un número importante de herramientas propietarias y de código abierto.

Por otro lado, existe una necesidad importante de permitir la interacción con sistemas y aplicaciones remotas, como por ejemplo aplicaciones móviles, por ello la

definición de las Interfaces de Programación de Aplicaciones (Application Programming Interface - API) toman un rol importante en el diseño e implementación dado que permiten interacciones con sistemas o aplicaciones existentes o nuevas heterogéneas.

Los WS, brindan un soporte adecuado y de manera transparente a la interacción de sistemas remotos. El consorcio W3C [6] define los WS como sistemas de software diseñados para soportar una interacción interoperable máquina a máquina sobre una red. Los WS suelen ser APIs Web que pueden ser accedidas dentro de una red (principalmente Internet) y son ejecutados en el sistema que los aloja.

Una clase de WS son los basados en llamada a procedimientos remotos (Remote Procedure Calls, RPC), los cuales son familiares a muchos desarrolladores. La interfaz es definida a través un archivo WSDL (Web Services Description Language WSDL) el cual es un formato XML. Las primeras herramientas para WS estaban centradas en esta visión y tuvo un auge importante. Sin embargo, ha sido algunas veces criticado por no ser débilmente acoplado, ya que suele ser implementado por medio del mapeo de servicios directamente a funciones específicas del lenguaje o llamadas a métodos.

Los WS pueden ser implementados siguiendo los conceptos de la Arquitectura Orientada a Servicios (Service-Oriented Architecture, SOA), donde la unidad básica de comunicación es el mensaje. Esto es típicamente referenciado como servicios orientados a mensajes. Los WS basados en SOA son soportados por una gran cantidad de desarrolladores de software y analistas. Al contrario que los Servicios Web basados en RPC, este estilo es débilmente acoplado, lo cual es preferible ya que se centra en el “contrato” proporcionado por el documento WSDL, más que en los detalles de implementación subyacentes.

En los últimos años se ha popularizado un estilo de arquitectura de Software conocido como REST (Representational State Transfer). Este nuevo estilo ha supuesto una nueva opción de estilo de uso de los Servicios

Web. Los WS basados en REST intentan emular al protocolo HTTP o protocolos similares mediante la restricción de establecer la interfaz a un conjunto conocido de operaciones estándar (GET, PUT, POST, DELETE, etcétera), por lo tanto, este estilo se centra más en interactuar con recursos con estado, que con mensajes y operaciones. En particular, el lanzamiento de REST como protocolo de intercambio y manipulación de datos en los servicios de internet produjo un gran cambio en el desarrollo de software en los últimos años. Este nuevo enfoque de desarrollo de proyectos y servicios web fue definido por Roy Fielding, el padre de la especificación HTTP y uno de los referentes internacionales en todo lo relacionado con la Arquitectura de Redes [7].

En la actualidad existe una gran cantidad de proyectos o aplicaciones que disponen de una API REST para la creación de servicios profesionales. Twitter, YouTube, los sistemas de identificación con Facebook, y cientos de empresas generan negocio gracias a REST y las APIs REST. REST es una interfaz entre sistemas que usa HTTP para obtener datos o generar operaciones sobre esos datos en todos los formatos posibles, como XML, JSON, HTTP, etc. En los últimos años logró un gran impacto en la web que prácticamente logró desplazar a SOAP y las interfaces basadas en WSDL por tener un estilo bastante más simple de utilizar y sobre todo por su eficiencia [8].

El presente trabajo propone un mecanismo de generación de API REST a partir de versiones existentes de API Java, en el contexto de MDD, para la construcción de WS's. Aplicar esta técnica mediante la transformación de modelos se diferencia de otras formas convencionales, las cuales se basan en generar un Árbol de Sintaxis Abstracta (Abstract Syntax Tree AST) mediante algún parser de JAVA. Además, nuestra propuesta permite generar código hacia distintas implementaciones de WS REST a partir de un modelo JAVA.

2. LÍNEA DE INVESTIGACIÓN y DESARROLLO.

Para lograr la generación automática de API REST, aplicando transformación de modelos, es muy importante partir de un modelo origen bien definido.

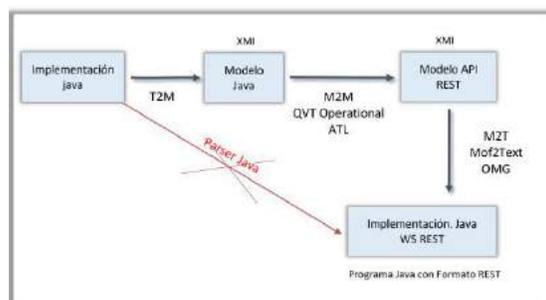


Figura 1: Etapas de transformación propuesta

En la figura 1, se muestra esquemáticamente nuestra propuesta, donde se define un proceso de transformaciones de la siguiente manera:

1. A partir de un programa java, con una interfaz bien definida, generar un modelo java (XMI) mediante una transformación texto a modelo (T2M).
2. Luego del modelo java (XMI), utilizando un lenguaje de transformación de modelo convencional (QVT Operacional, ATL; etc.), generar un modelo API REST (XMI) mediante una transformación de modelo a modelo (M2M).
3. Finalmente, del modelo API REST (XMI), generaremos el código java que implementará una API REST, a través de una transformación de modelo a texto (M2T).

Con la presente propuesta es sencillo diseñar aplicaciones que permitan la interacción entre sistemas heterogéneos tal como lo muestra la figura 2.

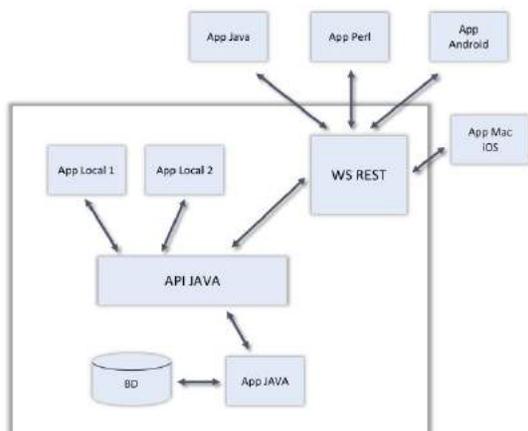


Figura 2: Diseño de un sistema utilizando API REST

En la actualidad existen trabajos que se vinculan. EMF-REST [9] es una herramienta que aprovecha las técnicas MDE para generar APIs Web RESTful a partir de los modelos EMF, promoviendo así la gestión de modelos en entornos distribuidos. La API Web RESTful generada se basa en bibliotecas y estándares bien conocidos con el fin de facilitar su comprensión y mantenibilidad. A diferencia de otros enfoques basados en el MDE para la generación de servicios web, EMF-REST proporciona un mapping directo para acceder a los modelos de datos mediante servicios Web siguiendo los principios de REST. Además, EMF-REST aprovecha características del modelo y de la Web, como el modelo Validación y seguridad, respectivamente.

En [10] los autores definen un mecanismo de diseño y descripción de API REST basado en formalismos de redes de Petri coloreadas, lo cual permite escalabilidad, extensibilidad e interoperabilidad. Además, permite una mayor facilidad para el testing por estar basado en un formalismo

3. RESULTADOS ESPERADOS

El proceso de transformaciones inicia con un programa escrito en el lenguaje JAVA, el cual es transformado a una representación de formato XMI, precisamente ECORE, conforme al metamodelo de dicho lenguaje de

programación. Este proceso es llevado a cabo utilizando la herramienta MoDisco.

Posteriormente, sobre dicho modelo se aplica una transformación Model-To-Model definida en QVT Operacional que nos permitirá generar una versión API REST del programa original. Finalmente, se obtiene el código JAVA que implementará la API REST a través de una transformación Model-To-Text definida con MOF2Text vía la herramienta Acceleo [11].

La contribución principal de este trabajo es proponer un mecanismo de transformación de programas JAVA, haciendo uso de estándares y herramientas existentes en el contexto de desarrollo dirigido por modelos, obteniendo como resultado una aplicación API REST que permite la interacción entre sistemas heterogéneos, promoviendo así la gestión de modelos en entornos distribuidos.

El mecanismo de generación de código es muy particular, dado que el mismo está basado en el tratamiento de los modelos JAVA (XMI), y no sobre el código fuente JAVA, dada la existencia de nuevas herramientas que permiten obtener dicho modelo, en caso contrario sería necesario manipular un parser del lenguaje.

4. FORMACION DE RECURSOS HUMANOS

Durante el desarrollo de esta línea de investigación, integrantes del grupo de trabajo están actualmente trabajando es su tesis de Licenciatura, Magister en Ingeniería de Software y Tesis de Doctorado.

También, se han formado ayudantes de segunda en las asignaturas de Análisis y Diseño de Sistemas, Ingeniería de Software, Base de Datos y Proyecto.

Los temas abordados en esta línea de investigación brindan un fuerte aporte al proceso de perfeccionamiento continuo de los autores de carreras de computación en Universidades Nacionales como del exterior.

5. BIBIOGRAFÍA

- [1] B. Selic, The pragmatics of model-driven development, IEEE Softw., vol. 20, no. 5, pp. 19-25, 2003.
- [2] Object Management Group, Object Management Group Std., Last access: May 2015. [Online]. Available: <http://www.omg.org>
- [3] J. Miller and J. Mukerji, Mda guide version 1.0.1, Object Management Group (OMG), Tech. Rep., 2003.
- [4] OMG, Meta Object Facility (MOF) 2.0 Query/View/Transformation Specification, Version 1.1, Object Management Group Std., Rev. 1.1, 2011. [Online]. Available: <http://www.omg.org/spec/QVT/1.1/>
- [5] OMG Unified Modeling Language (OMG UML), Superstructure, Version 2.4.1, Object Management Group Std., Rev. 2.4.1, 2011. [Online]. Available: <http://www.omg.org/spec/UML/2.4.1>
- [6] World Wide Web Consortium (W3C), World Wide Web Consortium (W3C) Std., último acceso: junio 2017. [Online]. Available: <http://www.w3.org/>
- [7] Architectural Styles and the Design of Network-based Software Architectures, UNIVERSITY OF CALIFORNIA, IRVINE Std., último acceso: junio 2017. [Online]. Available: <http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm>
- [8] L. Richardson and S. Ruby, Restful Web Services, 1st ed. O'Reilly, 2007.
- [9] H. Ed-Douibi, J. L. C. Izquierdo, A. Gómez, M. Tisi, and J. Cabot, "EMF-REST: generation of restful apis from models", CoRR, vol. abs/1504.03498, 2015. [Online]. Available: <http://arxiv.org/abs/1504.03498>
- [10] L. Li and W. Chou, "Design and describe rest api without violating rest: A petri net based approach." in ICWS. IEEE Computer Society, 2011, pp. 508-515. [Online]. Available: <http://dblp.uni-trier.de/db/conf/icws/icws2011.html#LiC11>
- [11] Acceleo, último acceso: mayo 2018. [Online]. Available: <https://www.eclipse.org/acceleo/>

Arquitectura de Software en el Proceso de Desarrollo Ágil. Una Perspectiva Basada en Requisitos Significantes para la Arquitectura.

Mg. Mirta E. Navarro¹, Mg. Marcelo P. Moreno², Lic. Juan Aranda³, Lic. Lorena Parra⁴, Lic. José R. Rueda⁵

Departamento de Informática - F.C.E.F. y N. - U.N.S.J.

Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan

¹mirtaenavarro@yahoo.com.ar²mpmoren@gmail.com³juanaranda@live.com⁴

⁵lorenaparra152@yahoo.com.ar⁵joserocardorueda@hotmail.com

Resumen

Recopilar, comprender y gestionar los requisitos es un aspecto crítico en todos los métodos de desarrollo. Esto también es cierto para las Metodologías Ágiles en la que la captura de requisitos es realizada en todo el proceso de desarrollo, con requisitos que van evolucionando y cambiando a lo largo del ciclo de vida. Este proceso es opuesto al enfoque de la Arquitectura del Software, donde los requerimientos deben ser identificados, recabados y comprendidos en las primeras etapas, ya que cambios posteriores afectan considerablemente el resultado final de la arquitectura. Este tratamiento con perspectivas diferentes, ha sido uno de los factores que ha causado la sensación de que las Metodologías Ágiles y la Arquitectura de Software van en direcciones diferentes y no pueden coexistir.

En este trabajo se presenta una línea de investigación en la que este equipo se ha enfocado en los últimos dos años y tiene como objetivo proponer un modelo que facilite la identificación y captura de los llamados “Requisitos Significantes para la Arquitectura”, permitiendo de esta manera la integración de aspectos arquitectónicos en el proceso de desarrollo de Sistemas de Información con Metodologías Ágiles, favoreciendo

de este modo, atributos de calidad y flexibilidad ante los cambios.

Palabras claves: *Software Architecture, Agile methodologies, Architecturally Significant Requirement, Information Systems*

Contexto

El presente trabajo se encuadra dentro del área de I/D de la IS y de los SI, y describe los lineamientos generales del proyecto de investigación: “*Arquitecturas de Software en el Proceso de Desarrollo Ágil, Una Perspectiva Basada en Requisitos*”, iniciado en enero de 2018, con una duración de dos años y que tiene como unidades ejecutoras al Departamento de Informática, FCEFyN de la UNSJ.

El grupo de investigación tiene una trayectoria de 16 años en diferentes proyectos vinculados a Metodologías de Desarrollo y Tecnologías, con numerosas publicaciones en diferentes ámbitos, y con la formación de recursos humanos en el área de interés.

Introducción

Las Metodologías Ágiles (MA) se centran en el trabajo en equipo, la adaptabilidad y colaboración dentro del equipo de desarrollo del software y

también entre los miembros del equipo y los usuarios finales. El uso de las MA, ha marcado una tendencia para su adopción al desarrollo de proyectos de software con necesidades cambiantes y a la espera de beneficios en el menor tiempo posible [1,2,3].

Desde la perspectiva de los requisitos con implicancias en aspectos arquitectónicos en particular, si bien las MA explican que éstos deben estar entre los de mayor prioridad y deben ser tenidos en cuenta desde el comienzo [4,5], pensamos que más que una falta de atención a la definición de los requisitos arquitectónicos al comienzo, la consideración de los mismos es un proceso no exento de dificultades y esto hace que típicamente se evite un trabajo inicial sustancial, en la suposición de que los requisitos siempre cambian, evolucionan y continúan cambiando a lo largo del ciclo de vida del proyecto. En ese sentido, generalmente se identifican y captan los requisitos, tanto funcionales como no funcionales, y luego se intenta definir una arquitectura de software que los cumpla.

La Arquitectura de Software (AS), en tanto, describe la solución de un sistema y por lo tanto tradicionalmente se piensa como una parte temprana de la fase de diseño [6] ya que reúne todos los requerimientos técnicos y operacionales y que son difíciles de cambiar durante el proceso de desarrollo. Es principalmente importante para satisfacer los requisitos no funcionales, que están relacionados a los atributos de calidad como el rendimiento, seguridad y escalabilidad.

Estas decisiones tempranas llevan un peso importante con respecto al desarrollo del resto de un sistema, condicionando decisiones posteriores que implican a su vez, otras ramificaciones subsiguientes. Esto supone una captura de requisitos sin cambios sustanciales en las etapas

intermedias y finales del desarrollo de un proyecto, visión que se contrapone a las dificultades para la identificación de requisitos arquitectónicos que se percibe en las MA.

Este tratamiento con enfoques diferentes en las etapas tempranas (y también en otros aspectos), ha sido uno de los factores que ha causado la sensación de que las MA y la AS van en direcciones diferentes y no pueden coexistir juntas [7]. Sin embargo, en los últimos años esta tendencia está cambiando [8,9], hasta el punto que varios autores hacen referencia al término “Arquitectura Ágil” (AA) [8,9]. En la AA, se enfatiza fuertemente en el concepto de los “Requisitos Significantes para la Arquitectura” (RSA). En tal sentido, aquí se propone la integración de las AS en el ciclo de vida de las MA, desde la perspectiva de la identificación de los RSA considerando diferentes propuestas existentes en el mismo sentido [9,10,11], que son la base para discutir y llevar a cabo más investigaciones sobre RSA. Pensamos que esos métodos y enfoques pueden ser complementarios entre sí, tomando lo mejor de cada uno de ellos, y a la vez podrían ser combinados con los hallazgos que han sido investigados por los autores de este artículo [12] en el área de las metodologías ágiles.

Requisitos de Importancia Arquitectónica.

Un Requisito Significante para la Arquitectura es un requisito que tendrá un efecto importante en la arquitectura, y que, si llegase a estar ausente, la arquitectura resultante será totalmente diferente. Tomando la definición de [10] los RSA “*son aquellos requisitos que tienen un impacto medible en una arquitectura de sistemas de software*”,

por lo que no sería posible diseñar una arquitectura adecuada, si no se identifican y capturan adecuadamente.

Un problema recurrente que se presenta con los RSA [8,9] es que generalmente toman la forma de requisitos no funcionales, pero en algunos casos también la forma de requisitos funcionales. Además, los RSA suelen ser subjetivos, relativos y también interactúan entre ellos. Son subjetivos, porque pueden ser vistos, interpretados y analizados de manera diferente por diferentes personas y en diferentes contextos; son relativos, porque la importancia de cada RSA se determina a veces a partir de su relación con otros RSA en un contexto dado, y se considera que interactúan entre ellos en el sentido de que al intentar alcanzar un RSA particular, puede a su vez afectar (en forma favorable o desfavorable) a otros RSA. Los RSA son más difíciles de comprender que los requisitos funcionales [10], por lo que en general pasan desapercibidos o no obtienen suficiente atención por adelantado, factor que es más acentuado cuando se utilizan MA, donde generalmente se expresan (muchas veces en forma contradictoria) de manera informal durante el análisis de requisitos. También son difíciles de validar cuando el proyecto está finalizado. Por todo lo expuesto, se considera que la identificación de los RSA, es una tarea difícil, que lleva mucho trabajo y es poco clara para los que no son expertos en arquitecturas [9]. Para abordar esta situación, algunos autores, han propuesto una serie de estrategias de identificación y captura de RSA a partir de documentos de requisitos tradicionales de una MA.

Enfoques.

Hay diferentes métodos para identificar y capturar los RSA. Algunos están basados en el conocimiento de los

objetivos del negocio [9], otros en un framework [10] y otros se basan en el conocimiento detallado del dominio [11]. El resumen de esos enfoques se detalla a continuación:

- El enfoque basado en un Framework propone una forma de identificar y capturar RSA a partir de una caracterización previa de los mismos sobre la base de un estudio empírico realizado con expertos.
- El enfoque basado en el Conocimiento del Dominio (CdD) en cambio, hace énfasis en las propiedades del medio ambiente y las suposiciones sobre él, más que en la caracterización de los RSA.
- Finalmente, un enfoque que tiene como propósito centrarse principalmente en los Objetivos de Negocio (ON) y posteriormente a partir de ellos, determinar cómo esos objetivos influyen en los requisitos de calidad que finalmente serán origen de muchos RSA.

Estos enfoques, más algunos otros que pudieran surgir en el periodo de investigación serán tenidos en cuenta en el proyecto y constituyen la base de la propuesta de investigación que se presenta en este trabajo.

Tópico de investigación, Desarrollo e Innovación

Los posibles beneficios de una integración de Arquitectura de Software en las Metodologías Ágiles, desde la perspectiva de la identificación y captura de los RSA, no es un tema lo suficientemente explorado en ambientes académicos, habiéndose encontrado escasas publicaciones relacionadas a los enfoques ya mencionados.

Actividades y Objetivos

Las actividades en las que se ha planificado el desarrollo de la investigación son las siguientes:

- Estudio de factores relevantes en la identificación y captura de los RSA.
- Analizar propuestas alternativas para la identificación y captura de los RSA.
- Fijación de criterios para la especificación de los RSA.
- Adopción de la MA para la integración con los RSA.
- Describir las características que se deben observar en el planteamiento y clasificación de requerimientos funcionales, no funcionales y RSA.

Todas estas actividades serán realizadas con el objetivo de proponer un Modelo de Identificación y Captura de los RSA, que pueda ser incorporado a los procesos de metodologías ágiles, respetando los requerimientos propios del dominio agilista. A partir de ese punto, se continuará con las siguientes actividades:

- Implementación de Sistemas de Información usando el Modelo de identificación/captura de RSA.
- Pruebas, validaciones y elaboración de conclusiones.
- Realización de transferencias al medio, divulgación científica y publicaciones

Formación de Recursos Humanos

El equipo de trabajo está compuesto por cinco docentes-investigadores, y cuatro alumnos adscriptos.

Con los resultados de la presente investigación, se harán actividades de

divulgación en publicaciones y presentaciones en eventos nacionales e internacionales y también asesorar trabajos de tesis de estudiantes de grado y/o posgrado.

En el periodo 2016-2017 se presentaron nueve trabajos en diversas publicaciones en Congresos Nacionales y Revistas Especializadas sobre la temática abordada. Además, se han asesorado los siguientes trabajos vinculados a la temática de referencia:

- Una tesis de maestría finalizada y aprobada en defensa pública (diciembre de 2017).
- Una tesis de grado finalizada y aprobada. (abril 2017).

Referencias

[1] Schwaber, K., Sutherland, J., "The scrum guide. The Definitive Guide to Scrum".En:

<https://www.scrumguides.org/docs/scrumguide/v2017/2017-Scrum-Guide-US.pdf>.
Accedido el 20/02/2017.

[2] Cohen, D., Lindvall, M., Costa, P., "An Introduction to Agile Methods" Pdf (2004) En:

http://www.cse.chalmers.se/~feldt/course/s/agile/cohen_2004_intro_to_agile_methods.pdf.
Accedido el 20/02/2018

[3] Breivold, H.P., Sundmark, D., Wallin, P. and Larsson, S., "What Does Research Say about Agile and Architecture?" en Proceedings of the Fifth International Conference on Software Engineering Advances (ICSEA), USA, (2010).

[4] Schwaber, K., Beedle M. "Agile Software Development with Scrum" 1st Edition. Prentice Hall PTR Upper Saddle River, NJ, USA 2001.

- [5] Kenneth Rubin “Essential Scrum: A Practical Guide to the Most Popular Agile Process”. 1st Edition. Addison-Wesley Signature Series. 2013.
- [6] Fowler, M. “Patterns of Enterprise Application Architecture” 1st Edition. Addison-Wesley. 2003.
- [7] Nord RL, Tomayko JE. “Software architecture-centric methods and agile development.” IEEE Software 2006; 23(2):47–53.
- [8] Babar MA, Brown AW, Mistrik I, “Agile Software Architecture Aligning Agile Processes and Software Architectures”. Morgan Kaufmann Publishers Inc. 2013.
- [9] Bass L, Clements P, Kazman R. “Software architecture in practice” 3rd ed. Addison-Wesley. 2012
- [10] Chen L, Babar MA, Nuseibeh B. “Characterizing architecturally significant requirements.” IEEE Software 2013; 30:38–45.
- [11] Azadeh Alebrahim, Maritta Heise. “Intertwining Relationship Between Requirements, Architecture, and Domain Knowledge”. ICSEA 2014: The Ninth International Conference on Software Engineering Advances.
- [12] Navarro M., Moreno M., Aranda J., et al. “Integración de Arquitectura de Software en el Ciclo de Vida de las Metodologías Ágiles.” Una Perspectiva Basada en Requisitos. XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017). (566-569). ISBN 978-987-42-5143-5.

Planificación de las Pruebas del Software

Gladys Kaplan^{1,2}, Walter Panessi¹, Claudia Ortiz¹, Eugenia Cespedes¹

¹ Departamento de Ciencias Básicas, Universidad Nacional de Luján (UNLu)

² Departamento de Ingeniería e Innovación Tecnológica, Universidad Nacional de La Matanza (UNLaM)

(gkaplan@unlam.edu.ar, wpanessi@unlu.edu.ar, cortiz@unlu.edu.ar, eugeniacespedes@outlook.com)

RESUMEN

La complejidad actual en la construcción de los sistemas de software, ha impuesto la necesidad de mejorar los procesos de construcción con el menor costo y esfuerzo posible. Las pruebas del software, a diferencia de otras etapas del proceso de construcción, no tienen la misma visibilidad para el usuario, lo que determina que ante presiones de tiempo o costos, sea esta etapa la que se redefina según lo planeado, a pesar de que esta decisión atenta directamente sobre la calidad del software. Para asegurar la correcta construcción del software, en relación a las pruebas, se hace indispensable pensarlas integradas y lo más automatizadas posible. En el proyecto de investigación, se ha trabajado sobre la derivación semiautomática de Casos de Prueba a partir de escenarios futuros y en el presente trabajo se ha avanzado sobre la planificación de esos Casos de Prueba obtenidos, para asegurar que la funcionalidad propuesta en los escenarios contemple las dependencias de recursos y funcionales existentes entre las tareas. De esta manera se utiliza la integración de los escenarios para identificar las

dependencias y ordenar la ejecución de los Casos de Prueba en las denominadas suites de prueba, que permiten analizar la funcionalidad parcial y total del futuro sistema de software.

Palabras clave: ingeniería de requisitos, casos de prueba funcionales, suite de pruebas.

CONTEXTO

Este trabajo se lleva a cabo en el marco del Proyecto de Investigación “Generación semi automática de casos de prueba a partir de escenarios”, radicado en la Universidad Nacional de Luján, en su Departamento de Ciencias Básicas. [DISPOSICION CDD-CB:027-15].

Como parte de las actividades se ha estudiado un mecanismo para generar tempranamente los casos de prueba partiendo del conocimiento obtenido en la etapa de Ingeniería de Requisitos (IR) [1], particularmente en el proceso de requisitos basado en escenarios [2].

Se profundizó luego la relación escenarios futuros – casos de prueba (EF-CP) y se describió un mecanismo de derivación semiautomática [3].

1. INTRODUCCIÓN

Desde hace mucho tiempo se plantea la necesidad de escribir los test de aceptación conjuntamente con los requisitos. Esto permite, por un lado, asegurarse que los requisitos sean verificables y por otro, que el futuro desarrollo y su respectivo desarrollador, cuente con ejemplos que le permitan comprender mejor el requisito.

Este trabajo es una evolución de otros anteriores, siempre basándonos en el proceso de requisitos [2], que crea un glosario denominado LEL [4], [5] del dominio de aplicación. El proceso continúa con la construcción de un conjunto de Escenarios Futuros (EF) a partir de los Escenarios Actuales (EA) [6] y su consecuente decisión de incluirlos en el nuevo sistema.

Trabajos previos de diferentes autores han planteado la necesidad de generar los casos de prueba a partir de varios artefactos que permiten documentar requisitos, como la generación de Casos de Prueba a partir de Casos de Uso [7], [8], [9], [10], [11], a partir de algoritmos metaheurísticos [12], utilizando el lenguaje de transformación QVT, otros a partir de diagramas de secuencia extendidos [13] o a partir de diagramas de actividad [14], [15], [16], [17], a partir de diagramas de secuencia [18] o de diagramas de estado [19]. Algunos lo han planteado en general para el lenguaje UML [20], [21], [22], [23] también a partir de especificación de Requisitos [24].

En el presente trabajo continuamos la idea de derivar en forma semiautomática los Casos de Prueba a partir de los Escenarios Futuros [3] pero se incorpora la necesidad de planificar las pruebas agrupando funcionalidad en diferentes niveles de pruebas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Como se ha mencionado anteriormente, los CP obtenidos desde los escenarios futuros, son estrictamente funcionales y con un alto grado de contextualización. Esto permite darle al caso de prueba información muy valiosa para comprender la mejor manera de analizar la funcionalidad propuesta. Es de destacar que los EF no son independientes entre sí, sino que se rigen por dependencias operativas, de recursos y de contextos. Este ordenamiento funcional se representa en el proceso de requisitos al modelar los escenarios integradores [25]. Cuando los EF son integrados se construyen las jerarquías de escenarios que luego se ordenan en secuencias para su ejecución, conformando los escenarios integradores. El mismo proceso de jerarquización y secuencia se aplicó a los casos de prueba para planificar las pruebas del software.

Las mismas pruebas del software se repiten cuando hay cambios en los requisitos. Existen requisitos que no se modifican, lo que hace que parte de la funcionalidad se mantenga inalterable durante parte o todo el proceso de construcción. Debido al costo de volver a realizar todas las pruebas, es que se

permite seleccionar qué parte de la funcionalidad se desea volver a probar. Esto requiere de la intervención del ingeniero de software y de una aplicación que controle y alerte de las dependencias con otros requisitos. Como puede verse en la Tabla 2 la primera columna es para identificar la funcionalidad con el nombre del escenario futuro, para ello se incorpora el escenario raíz de cada jerarquía obtenida en la integración. La columna de Última Prueba cumple la función de informar si después del último cambio realizado a ese escenario la prueba fue Aprobada o si se modificó el escenario y no se realizó o fue rechazada. Obviamente las pruebas rechazadas o que no se hicieron deben rehacerse obligatoriamente, pero cuando existen varios escenarios que no fueron modificados es decisión del ingeniero de software volver a probar todo o no hacerlo. Esto atiende a pruebas de gran volumen con parte de los requisitos estables.

Escenario	¿Se debe probar solo?	Última Prueba
EFa	No	Aprobado
EFb	SI	
EFc	SI	
EFd	SI	
Efe	SI	
EFf	Si	

Tabla 2 – Tabla de alcance de la prueba

En un primer momento del proceso se determinó cuáles eran los CP para cada

EF. En la Fig.1 se da un ejemplo aclaratorio del proceso:

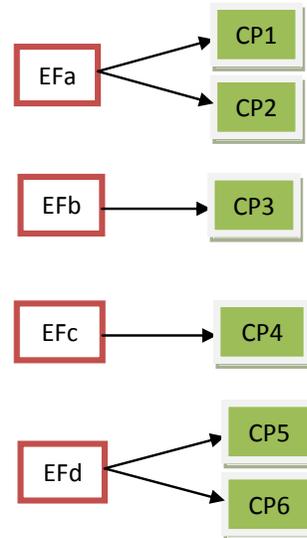


Fig. 1 – Derivación de los CP

Con los CP identificados para cada EF, se debe ir a la integración para generar las suites de prueba. Tomar cada integrador y ordenar los CP correspondientes.

Escenarios Integradores	Escenarios Futuros y sus CP	
I1	EFe	CP7
	EFa	CP1 y CP2
	EFc	CP4
I2	EFb	CP3
	EFd	CP5 y CP6

Tabla 3 – Escenarios integradores con sus EF y cada EF con sus CP derivados

Con esta información se está en condiciones de crear las suites de pruebas. Cada escenario integrador corresponde a una suite. Por lo tanto, según el ejemplo, tenemos 2 Suite de Pruebas con la siguiente secuencia:

Suite 1 {CP7, CP1, CP2, CP4}

Suite 2 {CP3, CP5, CP6}

En resumen, por un lado se cuenta con las pruebas individuales que corresponde a como se debe probar cada EF. Por el otro lado, tenemos los conjuntos de escenarios que se agrupan en integradores que a su vez determinan las suites de pruebas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados han sido alentadores debido a su vinculación directa con el proceso de requisitos basado en escenarios que ha sido ampliamente probado en más de 100 casos. De todas maneras se planifica probar todo el mecanismo propuesto en nuevos casos reales para mejorar su completitud y refinar el proceso. También se espera analizar cómo se incorporan los cambios de los requisitos en las iteraciones de la generación de los CP.

Como se aclaró en la introducción, el presente trabajo se ha realizado con base en el proceso de requisitos basado en escenarios que modela la solución con escenarios futuros, pero se considera que el proceso es viable para Casos de Uso con algunas pocas modificaciones.

4. FORMACIÓN DE RECURSOS HUMANOS

Se planifica la finalización de la carrera de grado Licenciatura en Sistemas de Información de la alumna Eugenia Céspedes (31 materias aprobadas) y del alumno Julián Massolo (34 materias

aprobadas) en UNLu. También la presentación de la tesis de Maestría en Ingeniería de Software de Claudia Ortiz y Walter Panessi en UNLP y la finalización de la tesis doctoral de Gladys Kaplan y de David Petrocelli en UNLP.

5. BIBLIOGRAFÍA

- [1] Kaplan Gladys, Doorn Jorge, Panessi Walter, Ortiz Claudia, Céspedes Eugenia, Massolo Julian, Petrocelli David, "Generación semi automática de casos de prueba a partir de escenarios", WICC 2015.
- [2] Leite, J.C.S.P., Doorn, J.H., "Perspectives on Software Requirements: An introduction" en el libro "Perspectives on Software Requirements", Kluwer Academic Publishers, EEUU, ISBN: 1-4020-7625-8, Capítulo 1, 2004.
- [3] Kaplan Gladys, Doorn Jorge, Panessi Walter, Ortiz Claudia, Céspedes Eugenia, "Derivación de casos de prueba a partir de escenarios", WICC 2017
- [4] Leite J.C.S.P., Franco, A.P.M., (1990) "O Uso de Hipertexto na Elicitação de Linguagens da Aplicação", Anais de IV Simpósio Brasileiro de Engenharia de Software, SBC.
- [5] Hadad G.D.S., Doorn J.H., Kaplan G.N. Creating Software System Context Glossaries, In: Mehdi Khosrow-Pour (ed) Encyclopedia of Information Science and Technology. IGI Global, Information Science Reference, Hershey, PA, USA, ISBN: 978-1-60566-026-4, 2nd edn, Vol. II. 2008.
- [6] Leite, J.C.S.P., Rossi, G., Balaguer, F., Maiorana, V., Kaplan, G., Hadad, G., Oliveros, A., "Enhancing a Requirements Baseline with Scenarios", Requirements Engineering Journal, Vol.2, N° 4, 1997.
- [7] M. Riebisch, I. Philippow, and M. Götze, "UML-Based Statistical Test Case Generation," in Revised Papers from the International Conference NetObjectDays on Objects, Components, Architectures, Services, and Applications for a Networked World, 2003, pp. 394-411.
- [8] Palacio Liliana González, "Método para generar casos de prueba funcional en el desarrollo de software", Revista Ingenierías Universidad de Medellín, vol. 8, No. 15 especial, pp. 29-36 - ISSN 1692-3324, 150 p. Medellín, Colombia, 2009

- [9] Jordán Enriquez, Odalys; Vázquez Ruiz, Orelvis, "Generacion de Casos de Prueba a partir de Casos de Uso en las pruebas del software", Ingeniería Industrial, vol. XXVII, núm. 1, 2006, pp. 7-10 Instituto Superior Politécnico José Antonio Echeverría La Habana, Cuba ISSN: 0258-5960, 2006.
- [10] Natalia Correa, Roxana Giandini, "Casos de Prueba del Sistema Generados en el Contexto MDD/MDT", LIFIA- Laboratorio de Investigación y Formación en Informática Avanzada, Universidad Nacional de La Plata, 40 JAIIO, ASSE 2011.
- [11] Javier J. Gutiérrez, María J. Escalona, Manuel Mejías y Jesús Torres, Hacia una propuesta de pruebas tempranas del sistema, XV Jornadas de Ingeniería del Software y Bases de Datos. JISBD 2006.
- [12] Raquel Blanco, Eugenia Díaz, Javier Tuya," Generación automática de casos de prueba mediante búsqueda dispersa", Revista Española de Innovación, Calidad e Ingeniería del Software, Vol.2, No. 1, 2006.
- [13] Beatriz Pérez Lamancha, Macario Polo, "Generación automática de casos de prueba para líneas de producto de software", Revista Española de Innovación, Calidad e Ingeniería del Software, Vol.5, No. 2, 2009.
- [14] Chen Mingsong, Qiu Xiaokang, Li Xuandong, "Automatic Test Case Generation for UML Activity Diagrams", Journal of Object Technology. Vol. 8, No. 3, May/June 2009.
- [15] Wang Linzhang, Yuan Jiesong ; Yu Xiaofeng ; Hu Jun ; Li Xuandong; Zheng Guoliang, "Generating test cases from UML activity diagram based on Gray-box method", Software Engineering Conference, 2004. 11th Asia-Pacific
- [16] P. N. Boghdady, Nagwa L. Badr, M. A. Hashim, Mohamed F. Tolba, "Test Cases Automatic Generator (TCAG): A Prototype", First International Conference, AMLTA 2012, Cairo, Egypt, December 8-10, 2012.
- [17] Boghdady, P.N. Badr, N.L.; Hashim, M.A.; Tolba, M.F., "An enhanced test case generation technique based on activity diagrams", Computer Engineering & Systems (ICCES), 2011
- [18] Javed, A.Z. "Automated Generation of Test Cases Using Model-Driven Architecture", Automation of Software Test, 2007.
- [19] Mourad Badri, Linda Badri and Maxime Bourque-Fortin, "Automated State-Based Unit Testing for Aspect Oriented Programs A Supporting Framework", Journal of Object Technology. Vol. 8, No. 3, May/June 2009.
- [20] Abbors, F., Backlund, A.; Truscan, D., "MATERA - An Integrated Framework for Model-Based Testing", Engineering of Computer Based Systems (ECBS), 17th IEEE International Conference 2010
- [21] Natalia Correa, Roxana Giandini (2012). Casos de Prueba del Sistema Generados en el Contexto MDD/MDT. 41 JAIIO - ASSE 2012 - ISSN: 1850-2792 - Pág. 91-105
- [22] Marc-Florian Wendland, Ina Schieferdecker and Alain Vouffo-Feudjio. Requirements-driven testing with behavior trees. Fourth International Conference on Software Testing, Verification and Validation Workshops. 2011.
- [23] Yasmine Ibrahim Salem y Riham Hassan. 2011. Requirement-Based Test Case Generation and Prioritization. 978-1- 61284-185-4/111 - 2011 IEEE
- [24] Bill Hasling, Helmut Goetz, Klaus Beetz. Model Based Testing of System Requirements using UML Use Case Models. International Conference on Software Testing, Verification, and Validation. 2008
- [25] Gladys Kaplan, Graciela Hadad, Jorge Doorn, "Apunte de Ingeniería de Requisitos", UNLaM

Evaluación de Variantes de Inspección en la Ingeniería de Requisitos

Alberto Sebastián¹, Graciela D. S. Hadad^{1,2}, Tomás Damonte¹, Ariel Vera², Ezequiel Robledo², Daniela Raffo¹

¹Facultad de Ingeniería y Tecnología Informática, Universidad de Belgrano

²Escuela de Informática, Universidad Nacional del Oeste

{alberto.sebastian, graciela.hadad, tomas.damonte}@comunidad.ub.edu.ar, avera@uno.edu.ar, ezeroble@hotmail.com.ar, daniela.raffo@comunidad.ub.edu.ar

RESUMEN

Es frecuente durante el proceso de requisitos construir modelos escritos en lenguaje natural, pues facilitan la elicitación, validación y negociación con el cliente. Estas facilidades se incrementan cuando los modelos se escriben usando el lenguaje propio del cliente. Para ello, se suele construir tempranamente un glosario con los términos utilizados en el contexto, denominado Léxico Extendido del Lenguaje. Este permite mejorar la comunicación entre los involucrados, sirviendo de apoyo a la descripción de modelos subsecuentes. Aunque, como todo modelo generado siguiendo heurísticas, presenta habitualmente inconsistencias, errores, omisiones y ambigüedades. Estas últimas aparecen básicamente por el uso del lenguaje natural. Estudios de estimación de completitud sobre este modelo establecieron la ocurrencia de un número significativo de omisiones. Es relevante poder detectar estos defectos y corregirlos oportunamente, evitando su propagación sobre otros modelos. Se han diseñado variantes de la técnica de inspección sobre el modelo léxico, las cuales identifican distintos tipos de defectos aunque su eficiencia y eficacia no son fácilmente comparables dado que no apuntan a detectar los mismos defectos. Se propone un análisis de estas variantes que permita establecer cuál es la más apropiada según el tipo de defecto que se requiera atender y el tiempo disponible.

Palabras clave: Ingeniería de Requisitos, Inspección de Modelos, Completitud de

Modelos, Ambigüedad.

CONTEXTO

La propuesta que se presenta es parte de los proyectos de investigación “Gestión de la calidad de un modelo léxico en el proceso de requisitos” de la Universidad de Belgrano y “Tratamiento de los factores situacionales y la completitud en la ingeniería de requisitos” de la Universidad Nacional del Oeste.

1. INTRODUCCIÓN

El proceso de requisitos involucra comprender el comportamiento actual en un contexto y definir la situación futura al incorporar un sistema de software [1]. En este proceso suelen utilizarse modelos escritos en lenguaje natural dado su cercanía a los clientes [2], aún cuando este tipo de modelos presentan defectos, principalmente del tipo ambigüedades y omisiones [3, 4, 5, 6].

Se han realizado varios estudios sobre la completitud de modelos en lenguaje natural en la Ingeniería de Requisitos, arrojando niveles excesivamente altos de omisiones [7, 8, 9]. Otros estudios se centraron en la ambigüedad de estos modelos, tal es el trabajo de Ben Achour et al. [10], quienes a través de un estudio empírico observaron que el 50% de los casos de uso contenían errores de terminología.

Dado que estos modelos construidos en el proceso de requisitos son la base para etapas posteriores del desarrollo de software, es necesario que presenten un nivel de calidad adecuado, pues los defectos no identificados tempranamente se trasladarán en cascada a modelos posteriores. Un mecanismo para

reducir estos defectos es aplicando técnicas de verificación, tales como las revisiones y las inspecciones [11].

La inspección de software [12, 13] fue diseñada como un proceso formal de detección de defectos en código fuente, posteriormente adaptada a modelos. Este proceso se compone de seis pasos bien definidos y con roles específicos en cada paso: inspector, productor (autor), moderador y escriba. Los pasos del proceso son:

1. Planeamiento: identificar el material a inspeccionar y las personas a cumplir con cada rol, y establecer la fecha de reunión.
2. Apreciación global: transmitir a los inspectores una descripción general del material a inspeccionar.
3. Preparación: leer el material por parte de los inspectores.
4. Reunión: determinar cuáles son defectos en el material por parte de los inspectores y transmitirlos a los productores del mismo, asistidos por el moderador mientras el escriba registra los hechos.
5. Corrección: eliminar los defectos del material revisado, a cargo de los productores.
6. Seguimiento: determinar el estado de la corrección, a cargo del moderador.

En general, en este proceso formal de revisión los defectos son catalogados por tipo y por severidad, y se cuantifican los tiempos de detección, para mejorar la eficiencia y eficacia del proceso mismo.

Las inspecciones han surgido como una de las técnicas de aseguramiento de calidad más eficaces en la ingeniería de software, debido a lo cual se han diseñado diferentes variantes aplicadas a modelos elaborados en el proceso de requisitos [14, 15, 16]. Estas variantes se han clasificado según el modo de detección de defectos en la etapa de preparación [17] en: lectura ad-hoc, lectura usando checklist, lectura usando procedimientos y lectura constructiva.

La lectura ad-hoc se hace en una modalidad desestructurada, donde el inspector recibe muy poco apoyo para encontrar defectos. Esto no significa que los participantes de la inspección no escruten el

artefacto sistemáticamente. Sólo significa que ningún plan previo está disponible y todo se deja librado a la experiencia del inspector.

La lectura con checklist le da un apoyo más fuerte al inspector en la forma de una lista de preguntas y controles que tienen que ser contestados y revisados en un cierto orden preestablecido. Este enfoque tiene algunas debilidades, como la falta de ayuda en comprender el artefacto bajo estudio y pobre adaptación a un ambiente de desarrollo dado.

La lectura basada en procedimientos da una guía detallada al inspector sobre cómo encontrar defectos específicos, independizándose de la experiencia del inspector.

La lectura constructiva va más allá que meramente revisar un artefacto y producir una lista de defectos, pues durante la lectura el inspector produce una nueva representación del material bajo estudio, la que se analiza posteriormente para detectar defectos.

En pocos trabajos se reportan evaluaciones sobre qué variante de inspección puede ser más provechosa, y si ellas dependen del modelo a verificar y del tipo de defectos que presentan. En algunos experimentos, se siguió la experiencia del inspector [15, 18] y los resultados mostraron que la importancia de la definición del procedimiento disminuía a medida que aumentaba la experiencia del inspector, y que inspectores novicios podían adquirir rápidamente conocimiento sobre defectos típicos y aspectos de calidad. Paech et al. [18] comprobaron que se detectaban en promedio más defectos usando la técnica de lectura basada en procedimientos frente a la lectura con checklist, aunque acarrea más tiempo de inspección.

2. LÍNEAS DE INVESTIGACIÓN E DESARROLLO

En un proceso particular de requisitos orientado al cliente [19, 20], donde se crean y utilizan modelos en lenguaje natural para lograr un mayor involucramiento de los clientes, el primer modelo que se construye es el Léxico Extendido del Lenguaje (LEL) [21], con el fin de obtener una comprensión acabada del vocabulario que se utiliza en el

contexto del problema. Este modelo se conforma de un conjunto de términos relevantes en dicho contexto, con sus definiciones, dadas por dos componentes: noción (denotación del término) e impacto (su connotación). El modelo considera la definición de términos sinónimos y homónimos, como también jerarquías de términos (términos genéricos y especializados) [20].

Este modelo debe tener la mayor calidad posible, dado que los restantes modelos harán uso de la terminología definida en él y algunos de ellos pueden derivarse del propio LEL [20]. Para lograr esta calidad, se han propuesto heurísticas y procedimientos para reducir defectos en este modelo.

Las heurísticas se enfocan principalmente en la creación del modelo LEL aportando guías de estilo y de contenido [6, 21, 22]. A pesar de la aplicación de estas heurísticas, el nivel de completitud estimado de este modelo ha sido muy bajo. En un estudio realizado por Doorn y Ridao [7] donde estimaron el tamaño de un modelo LEL a partir de nueve muestras del mismo, obtuvieron en el mejor de los casos un nivel de completitud del 51%. En un estudio posterior realizado por Litvak et al. [23], se confirmaron estos valores, aún cuando las muestras del modelo léxico habían sido construidas con heurísticas más precisas.

Por otro lado, se han diseñado mecanismos específicos para la verificación de este modelo. Los defectos a ser identificados han sido categorizados en: discrepancias (inconsistencias), errores, omisiones y ambigüedades, y calificados según su grado de severidad en: alto, medio y bajo [24].

En proyectos previos, se diseñó una técnica de inspección del modelo LEL basada en procedimientos [25] para guiar la detección de diversos tipos de defectos. En esta variante se completa un conjunto de formularios siguiendo un procedimiento para cada formulario, el cual detalla cómo completar el formulario y cómo identificar defectos a través de la información registrada. Es una variante que facilita la detección de defectos a inspectores novatos, aunque insume un tiempo considerable. Por otro lado, tal cual

fue propuesta en [25], no establece el grado de severidad de los defectos ni los tipifica; esto fue realizado posteriormente para facilitar evaluaciones de la misma. Las omisiones que permite detectar son relativamente simples y algunas requieren una gran cantidad de comparaciones semánticas, siendo las ambigüedades tratadas como una subclase de omisiones.

Posteriormente, se elaboró una lista de control (checklist) para utilizar como otra variante de la inspección del LEL. Esta lista de control fue refinada incrementalmente a medida que se probaba en diversos casos. En su última versión, contiene 44 ítems de control, y establece por cada ítem el tipo de defecto a identificar y su severidad.

Más recientemente, se diseñó otra variante de la inspección del LEL, basada en la lectura constructiva [24, 26]. Ésta genera un artefacto intermedio, que son mapas conceptuales, uno por cada término del LEL. Cada oración en la noción y en el impacto del término se representa como una proposición en el mapa conceptual del mismo. Los defectos se identifican a partir del análisis sistemático de cada mapa conceptual y de las relaciones entre mapas, mediante guías de análisis. Esta variante se focaliza en la detección de varios tipos de omisiones y ambigüedades con diverso grado de severidad, permitiendo incluso sugerir términos candidatos omitidos, aunque no se aboca a la detección de otros tipos de defectos. Aún cuando debe construirse un artefacto a partir del modelo LEL, se han obtenido resultados relativamente aceptables, en relación al consumo de tiempo y a la cantidad de defectos detectados [27]. Otra particularidad de esta variante es que sugiere correcciones al modelo LEL.

Cabe notar que estas tres variantes de inspección fueron desarrolladas en forma independiente una de otra, por lo que no todo defecto es detectado por todas las variantes, y algunos defectos son identificados por una sola de ellas, lo que dificulta una comparación efectiva de las variantes.

Se espera poder dar pautas a la gerencia de un proyecto de software sobre qué variante de

inspección puede ser más provechosa según el nivel esperado de calidad, los recursos humanos disponibles y los plazos a cumplir.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se ha llevado a cabo un primer análisis de cada variante respecto a qué defectos comunes identifican (ver Tabla 1), cuáles son identificados por pares de variantes y cuáles son exclusivos de una variante (ver Tabla 2).

Tabla 1. Defectos comunes a las variantes

Defecto	Tipo	Severidad
Términos sin referencia a otros términos	Omisión	Media
Términos no referenciados por ningún término	Omisión	Media
Referencia a término no identificada en la definición de un término	Omisión	Media
Término con más de un verbo principal en cada oración de noción e impacto	Omisión	Baja

Tabla 2. Defectos detectados solo por variante con mapas conceptuales

Defecto	Tipo	Severidad
Vocabulario mínimo redundante (con sinónimos internos)	Ambigüedad	Media
Uso de término del vocabulario mínimo en lugar de término del LEL	Ambigüedad	Media
Término del vocabulario mínimo no detectado como sinónimo del LEL	Omisión	Media
Frases omitidas en la noción o impacto del término	Omisión	Baja
Términos con oraciones subordinadas complejas	Ambigüedad	Alta
Omisión de términos de tipo Verbo (uso frecuente)	Omisión	Alta

Debe tenerse en consideración que la inspección con checklist contiene 44 ítems de control, la inspección basada en procedimientos requiere 10 formularios de detección, y la inspección basada en mapas conceptuales utiliza 13 pasos de detección (además de los pasos de construcción de los mapas). Las variantes con checklist y con formularios detectan principalmente errores y

omisiones, mientras que la basada en mapas conceptuales se aboca a ambigüedades y omisiones.

Se ha iniciado un experimento controlado para realizar comparaciones de las tres variantes utilizando modelos LEL generados independientemente para distintos casos de estudio, y considerando los tipos de defectos que detectan, su nivel de severidad, los tiempos que insumen y la experiencia de los inspectores.

Asimismo, se está desarrollando una herramienta de software, denominada Gestor de Inspecciones, implementada en una plataforma web, utilizando lenguaje PHP y base de datos relacional MySQL. La herramienta permite administrar los resultados de inspecciones, realizadas por ingenieros de requisitos autorizados, sobre modelos LEL, soportando el almacenamiento de resultados para las tres variantes de inspección. Esta herramienta facilitará el análisis de la eficiencia y efectividad de las distintas variantes de inspección.

Por otro lado, se espera establecer las causas de los defectos detectados, de manera tal de poder definir heurísticas que permitan evitar la ocurrencia de estos defectos, ya sea al elicitar conocimiento del contexto de aplicación o al modelar, es decir, heurísticas que se aboquen a completar información y a escribir el modelo LEL sin ambigüedades.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto de la Universidad de Belgrano participan tres investigadores, uno de ellos en formación y un alumno de la tecnicatura en Programación de Computadoras, cumpliendo su Trabajo Social Profesional, mientras que en el tema de completitud dentro del proyecto de la Universidad Nacional del Oeste participa un investigador con dos alumnos becarios.

5. BIBLIOGRAFÍA

- [1] Nuseibeh, B., Easterbrook, S. (2000) Requirements Engineering: A Roadmap. Future of SE Track 2000, pp.35-46
- [2] Ryan, K. (1993) The Role of Natural Language in Requirements Engineering, IEEE International

- Symposium on Requirements Engineering, San Diego, CA, pp. 240-242.
- [3] Zowghi, D., Gervasi, V. (2002). The Three Cs of Requirements: Consistency, Completeness, and Correctness, 8th International Workshop on Requirements Engineering: Foundation for Software Quality, Essen, Germany: Essener Informatik Beitiage, pp. 155-164
- [4] Berry, D.M., Kamsties, E. (2004) Ambiguity in Requirements Specification. En: Leite & Doorn (eds.) Perspectives on Software Requirements, pp.7-44. Kluwer Academic Publishers.
- [5] Hadad, G.D.S., Litvak, C.S., Doorn, J.H., Ridao, M.N. (2015) Dealing with Completeness in Requirements Engineering. En: Khosrow-Pour, M. (ed), Encyclopedia of Information Science and Technology, Third Edition, pp.2854-2863. IGI Global, Information Science Reference.
- [6] Litvak, C.S., Hadad, G.D.S., Doorn J.H. (2017) Nominalizations in Requirements Engineering Natural Language Models. En: Encyclopedia of Information Science and Technology, Fourth Edition, IGI Global.
- [7] Doorn, J.H., Ridao, M. (2003) Completitud de Glosarios: Un Estudio Experimental. 6th Workshop on Requirements Engineering.
- [8] Ridao, M., Doorn, J. (2006) Estimación de Completitud en Modelos de Requisitos Basados en Lenguaje Natural. 9th Workshop on Requirements Engineering, ISSN: 1413-9014, pp. 151-158.
- [9] Hadad, G.D.S., Litvak, C.S., Doorn, J.H. (2014) Problemas y Soluciones en la Completitud de Modelos en Lenguaje Natural. II Congreso Argentino de Ingeniería. ISBN:978-987-1662-51-7.
- [10] Ben Achour, C., Rolland, C., Maiden, N.A.M., Souveyet, C. (1999) Guiding Use Case Authoring: Results of an Empirical Study, International Symposium On Requirements Engineering (RE'99), Limerick, Irlanda, IEEE Computer Society Press, 1999, pp.36-43.
- [11] Porter, A.A., Votta, Jr., L.G., Basili, V.R. (1995) Comparing Detection Methods for Software Requirements Inspections: A Replicated Experiment, IEEE Transactions on Software Engineering, Vol. 21, N° 6, pp. 563-575.
- [12] Fagan, M.E. (1976) Design and Code Inspections to reduce Errors in Program Development, IBM Systems Journal, Vol.15, N° 3, pp. 182-211.
- [13] Fagan, M.E. (1986) Advances in Software Inspections, IEEE Transactions on Software Engineering, Vol.12, N° 7, pp. 744-751.
- [14] Fusaro, P., Lanubile, F., Visaggio, G. (1997) A Replicated Experiment to Assess Requirements Inspection Techniques, Empirical Software Engineering, Vol. 2, N° 1, pp.30-57.
- [15] Porter, A.A., Votta, Jr. (1998) Comparing Detection Methods for Software Requirements Inspections: A Replication Using Professional Subjects. Empirical Software Engineering, Vol. 3, N° 4, pp. 355-380.
- [16] Leite, J.C.S.P., Doorn, J.H., Hadad, G.D.S., Kaplan, G.N. (2005) Scenario Inspections. Requirements Engineering Journal, Springer-Verlag, Vol. 10, N° 1, pp. 1-21.
- [17] Regnell, B., Runeson, P., Thelin, T. (1999) Are the perspectives really different? Further experimentation on scenario-based reading of requirements. En: Requirements engineering with use cases - a basis for software development, Reporte Técnico 132, Lund University, pp.141-180.
- [18] Paech, B., Denger, C., Kerkow, D., von Kneten, A. (2005) Requirements Engineering for Technical Products: Integrating Specification, Validation and Change Management, capítulo X, Information Science Publishing, Maté & Silva (eds.), Londres, 2005, pp.153-169.
- [19] Leite, J.C.S.P., Doorn, J.H., Kaplan, G.N., Hadad, G.D.S., Ridao, M.N. (2004) Defining System Context using Scenarios. En: Perspectives on Software Requirements, Kluwer Academic Publishers. ISBN:1-4020-7625-8, cap.8, pp.169-199.
- [20] Hadad, G.D.S. (2018) Uso de Escenarios en la Derivación de Software. Tesis Doctoral, Universidad Nacional de La Plata.
- [21] Hadad, G.D.S., Doorn, J.H., Kaplan, G.N. (2009) Creating Software System Context Glossaries. En: Encyclopedia of Information Science and Technology, 2nd Edition, pp.789-794. IGI Global.
- [22] Litvak, C.S., Hadad, G.D.S., Doorn, J.H. (2014) Heurísticas para el modelado de requisitos escritos en lenguaje natural, CACIC 2014 - XX Congreso Argentino de Ciencias de la Computación, Buenos Aires, pp. 682-691.
- [23] Litvak, C.S., Hadad, G.D.S., Doorn, J.H. (2013) Mejoras semánticas para estimar la Completitud de Modelos en Lenguaje Natural, 1er. Congreso Nacional de Ingeniería Informática / Sistemas de Información, Córdoba.
- [24] Sebastián, A., Hadad, G.D.S., Robledo, E. (2017) Inspección centrada en Omisiones y Ambigüedades de un Modelo Léxico, CIBSE 2017 – XX Congreso Iberoamericano en Software Engineering, track Workshop on Requirements Engineering, Buenos Aires.
- [25] Kaplan, G., Hadad, G., Doorn, J. Leite, J. (2000) Inspección del Léxico Extendido del Lenguaje. III Workshop on Requirements Engineering, pp.70-91.
- [26] Sebastián, A., Hadad, G.D.S. (2016) Enhancing a Lexicon Model by Concept Mapping. En: Computer Science & Technology Series. XXI Argentine Congress of Computer Science – Selected Papers, G.E. Feierherd, P.M. Pesado & C.C. Russo (eds.), Editorial de la Universidad de La Plata (EDULP), pp. 139-151.
- [27] Sebastián, A., Hadad, G.D.S. (2016) Experimento Controlado en la Inspección de un Léxico mediante Mapas Conceptuales. III Congreso Argentino de Ingeniería, Resistencia, pp. 2738-2752.

MODELOS DE MADUREZ PARA LA MEJORA DE CALIDAD DE LOS INDICADORES DE DESARROLLO SOSTENIBLE

Ignacio Marcovecchio^{1,2}, Elsa Estevez^{1,3}, Pablo Fillotrani^{1,4}

¹ Depto. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur, Argentina

² United Nations University Institute on Computing and Society (UNU-CS), Macao SAR, China

³ Instituto de Ciencias e Ingeniería de la Computación, (UNS-CONICET), Argentina

⁴ Comisión de Investigaciones Científicas Provincia de Buenos Aires, Argentina

ignaciomarcovecchio@gmail.com, {ece,prf}@cs.uns.edu.ar

RESUMEN

Disponer de datos útiles y confiables para poder tomar decisiones informadas es uno de los requisitos necesarios para alcanzar los objetivos de la Agenda de Desarrollo de 2030. Sólo mediante mediciones que reflejen claramente la realidad se pueden desarrollar estrategias y asignar recursos que lleven a alcanzar los objetivos que buscan erradicar la pobreza, proteger el planeta y asegurar que todas las personas disfruten de paz y prosperidad antes del año 2030. El trabajo de investigación que se presenta en este documento tiene como objetivo contribuir con el cumplimiento de los objetivos de desarrollo sostenible a través de intervenciones que contribuyan a fortalecer las capacidades de las entidades responsables de producir datos que describan la realidad de los distintos países. En particular, esta investigación propone la utilización de Modelos de Madurez de la Capacidad (CMMs) como herramienta para fortalecer las instituciones y mejorar los procesos utilizados para la producción de indicadores sociales a nivel nacional. Cuanto más maduras sean las instituciones dentro del ecosistema nacional de datos, mejor y más confiables serán los resultados que producen y, por lo tanto, mejores podrán ser las decisiones que permitan lograr el desarrollo sostenible.

Palabras clave: Objetivos de Desarrollo Sostenible, Modelos de Madurez de la Capacidad, Capacidad Institucional

CONTEXTO

El presente trabajo de investigación se realiza dentro de la colaboración entre la Universidad Nacional del Sur (UNS) y el Instituto de Computación y Sociedad de la Universidad de Naciones Unidas (UNU-CS). En UNU-CS, este trabajo representa una de las cuatro líneas de investigación del proyecto Data & Sustainable Development [1], que tiene entre sus objetivos mejorar la confianza en los datos de los indicadores. En la UNS, este trabajo se enmarca dentro de las líneas de investigación en Modelos y Aplicaciones de Interoperabilidad Semántica en Gobernabilidad Electrónica del Laboratorio de Investigación y Desarrollo en Ingeniería de Software y Sistemas de Información [2] y de Ingeniería de Software y Gobierno Digital del Instituto de Ciencias e Ingeniería de la Computación [3].

1. INTRODUCCIÓN

En septiembre de 2015 los líderes de 193 países definieron 17 objetivos para el desarrollo sostenible global. Estos objetivos, que se conocen como los Objetivos para el Desarrollo Sostenible (Sustainable Development Goals – SDGs), definen la agenda mundial de desarrollo y plantean un conjunto ambicioso de objetivos que deben mantener el balance entre los tres pilares fundamentales del desarrollo sostenible: inclusión social, desarrollo económico y sostenibilidad del medio ambiente. Los 17 objetivos intentan alcanzar 169 metas que van a ser controladas y evaluadas a través de 232 indicadores. La Comisión de Estadísticas de Naciones Unidas [4] desarrolló un sistema global de indicadores que permite controlar el progreso de los países para cumplir con los SDGs.

Un aspecto importante de la agenda de los SDGs es poder controlar el progreso en el cumplimiento de las metas, como también el desarrollo de herramientas y plataformas que sirvan de soporte a las actividades realizadas por los diferentes actores [5]. Se prevé que el control de los indicadores de los SDGs va a demandar grandes esfuerzos para que se produzcan datos confiables y de calidad, manteniendo la premisa de que “nadie se quede afuera” [6]. Sin embargo, estos objetivos representan un gran desafío a la capacidad de muchos países para medir el progreso para alcanzar las metas de los SDGs [7] ya que la capacidad de los actores principales del ecosistema debe ser potenciada para que se pueda utilizar y sacar provecho de los datos. Por este motivo, resulta indispensable asegurar que todos los países cuenten con un sistema nacional de estadísticas capaz de producir y controlar datos estadísticos confiables que cumplan con los estándares y expectativas globales [6].

Disponer de datos confiables es un factor crítico para poder transformar los SDGs en herramientas útiles para la toma de decisiones y la solución de los problemas. Sin datos actualizados y confiables, el diseño y la implementación de políticas adecuadas resulta muy difícil. Por estas razones, los datos juegan un rol fundamental en el sistema de control de los SDGs. Resulta indispensable poder disponer de datos de alta calidad que se puedan transformar en información que represente el progreso y que sirvan para poder decidir sobre la distribución de recursos, informar en la definición de políticas, y evaluar el impacto de los esfuerzos realizados para lograr los objetivos de la agenda.

Las Oficinas Nacionales de Estadísticas (National Statistical Offices – NSOs), que han sido históricamente las responsables de custodiar los datos para el bien público, siguen cumpliendo un rol clave en los esfuerzos de los gobiernos para producir datos y controlar el progreso en el cumplimiento de los SDGs. Para poder desarrollar este rol, sin embargo, deben ser capaces de cambiar y adaptarse de manera más rápida que en el pasado. Para ello, deben abandonar procesos de producción costosos e ineficientes, incorporar nuevas fuentes de datos y asegurar que los ciclos de datos están alineados con los ciclos de toma de decisiones. Uno de los

desafíos radica en que muchas NSOs aún no cuentan con la suficiente capacidad y financiamiento y, por lo tanto, resultan vulnerables a los intereses políticos y de otros grupos de influencia. Para proteger y mejorar la calidad de los datos se debe fortalecer a las NSOs y asegurar que puedan funcionar de manera autónoma e independiente de cualquier influencia política.

Considerando estos desafíos y la complejidad de los ecosistemas nacionales de datos, esta investigación busca fortalecer la capacidad de las instituciones estadísticas para la producción de datos útiles y confiables que permitan el control y la medición del progreso en el cumplimiento de los objetivos de los SDGs

El resto del trabajo está organizado como se detalla a continuación. En la sección 2 se describen los temas de investigación que cubre este trabajo a partir de la formulación de los objetivos específicos de investigación que se intentan alcanzar y la metodología de investigación que se definió para poder alcanzarlos. En la sección 3 se presentan los resultados obtenidos hasta el momento y los próximos pasos a ejecutar para alcanzar el resto de los objetivos planteados. Finalmente, en la sección 4 se explica la contribución esperada en cuanto a formación de recursos humanos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente trabajo tiene por objetivo general contribuir con el cumplimiento de los objetivos de la agenda de desarrollo sostenible para el año 2030 a través de intervenciones que contribuyan a fortalecer las capacidades de las entidades responsables de producir datos para los indicadores de los SDGs en los distintos países. Existe dentro de la comunidad internacional la necesidad de contar con datos confiables y hay numerosos esfuerzos para mejorar la calidad y la precisión de los datos estadísticos. Sin embargo, a partir de un análisis extensivo de las soluciones existentes [8], observamos dos limitaciones. La primera limitación es que gran parte de los trabajos en el área se enfocan en evaluar y mejorar la calidad de los datos producidos, prescindiendo de cómo fueron producidos. En esta investigación creemos que por lo que

representan los datos, la forma en la que los son producidos es más importante que los datos en sí mismos. La segunda limitación es que ninguno de dichos trabajos se centra en los datos y las demandas de la agenda de desarrollo sostenible. Debido a su relevancia y a las particularidades que la diferencian de la producción de datos estadísticos en general, esfuerzos en investigación y desarrollo que soporten la producción de datos con el fin específico de informar el estado de los indicadores de los SDGs son indispensables.

Para cumplir con el objetivo general de este trabajo y teniendo en cuenta el contexto en el cual se desarrolla, los siguientes son los objetivos específicos de investigación definidos:

- O1. Identificar y evaluar los mecanismos utilizados actualmente por una NSO para recolectar y reportar datos estadísticos relacionados con los SDGs.
- O2. Diseñar modelos y herramientas computacionales que permitan mejorar las deficiencias identificadas y facilitar la integración de datos heterogéneos de distintas fuentes.
- O3. En base a la evaluación de las herramientas actuales y las herramientas definidas, especificar un modelo de madurez para la producción de estadísticas que permita a las NSOs avanzar en la producción confiable de información de calidad para medir el cumplimiento de las metas de los SDGs.

La madurez es un reflejo del nivel de desarrollo organizacional que puede ser utilizado para determinar la capacidad de las organizaciones para desarrollar ciertas actividades. Bajo la hipótesis de que cuanto más maduras son las instituciones, mejor es la calidad¹ de los resultados que producen, este trabajo busca mejorar la capacidad de las entidades responsables de producir datos para el control de los indicadores de los SDGs mediante un mecanismo de evaluación de la madurez que no sólo describa la situación actual, sino que además permita determinar las acciones necesarias para poder mejorarla. Para tal fin, esta investigación propone la formulación de un nuevo CMM

prescriptivo y multidimensional para evaluar y mejorar la capacidad de las NSOs para reportar el progreso en el alcance de los SDGs a nivel nacional.

Los modelos de madurez son herramientas útiles para evaluar la calidad y la efectividad de los procesos. Los modelos de madurez pueden ser utilizados para identificar las fortalezas y debilidades organizacionales, y como herramientas para evaluaciones comparativas [9]. Los modelos de madurez prescriptivos superan a los modelos descriptivos ya que son útiles no sólo para evaluar la situación actual sino también para desarrollar mapas de ruta para implementar mejoras [10]. El modelo propuesto será multidimensional ya que para cada fase del proceso de producción de datos se definirán varias dimensiones de análisis.

2.1. Metodología

Este trabajo se encuadra dentro de las metodologías de modelos, que se centran en definir modelos abstractos de un sistema real; pero involucran además aspectos de las metodologías de procesos, que se utilizan para comprender y diseñar procesos que se desarrollan para realizar tareas [11].

Debido a que la motivación es la de contribuir al ecosistema mediante la introducción de nuevos artefactos y procesos para la construcción de dichos artefactos, este trabajo adopta una filosofía *Pragmática* [12] y sigue el enfoque de la *Ciencia del Diseño* (Design Science – DS). En particular, se opta por la vista de tres ciclos de la DS propuesta por Hevner, que entiende a la DS como tres ciclos de actividades fuertemente relacionados [13]. De acuerdo a esta vista de la DS, el reconocimiento de los tres ciclos – relevancia, diseño y rigor – en los proyectos de investigación posiciona y diferencia a la DS de cualquier otro tipo de paradigmas.

En el *ciclo de relevancia* es donde se vincula el contexto del proyecto de investigación con las actividades de la DS. En esta investigación el contexto es el ecosistema nacional de datos para los SDGs y es donde se identifican las deficiencias, los obstáculos y las oportunidades.

¹ La calidad de los datos en este contexto está definida por su confiabilidad y su utilidad para la planificación y la toma de decisiones

Adicionalmente, el mismo contexto es el que provee los criterios para la evaluación de los resultados. El *ciclo de diseño* es donde se construyen y evalúan los artefactos. Para guiar el desarrollo dentro del ciclo de diseño, para este proyecto se ha seleccionado un marco de trabajo que incluye un proceso de seis fases para el desarrollo de modelos de madurez [10]. Además del proceso, el marco de trabajo propone las distintas técnicas y procedimientos a utilizar en cada una de las fases. Finalmente, el *ciclo de rigor* es donde se conectan las actividades de la DS con la base de conocimientos de fundamentos científicos, experiencias y experticia que informan al proyecto de investigación. Los resultados de la investigación, junto con las experiencias adquiridas, serán las contribuciones de esta investigación que agregan valor a dicha base de conocimientos.

La estrategia de investigación seleccionada es *Investigación-Acción* (Action Research) [14] y debido a que el ecosistema de datos será estudiado y evaluado de acuerdo a su evolución en el tiempo, esta investigación se enmarca dentro de un horizonte temporal *longitudinal*.

Todas las decisiones de diseño de la investigación mencionadas anteriormente están basadas en la visión de los autores en función de su comprensión del dominio y de los objetivos perseguidos, y están sustentadas por una extensa revisión de la bibliografía del dominio. Por ejemplo, un estudio que analizó más de 200 artículos en el área de modelos de madurez resalta que el enfoque de la DS es el más utilizados cuando se desarrollan nuevos modelos de madurez [15].

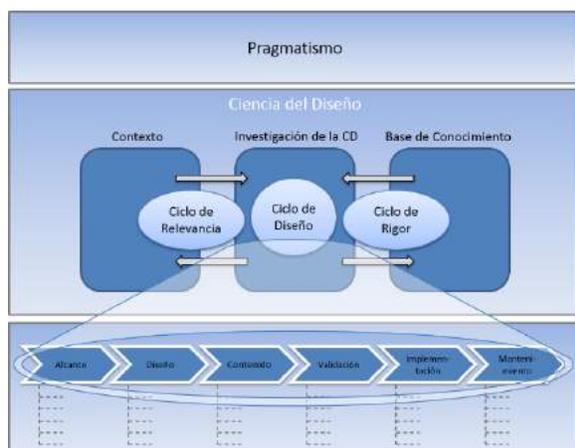


Figura 1: Diseño de la Investigación

Mientras que la Figura 1 ilustra y resume el diseño metodológico de esta investigación, a continuación se explica cómo las fases del proceso de desarrollo se instancian en este trabajo.

Para definir el *alcance* se definió el enfoque y se identificaron todos los actores involucrados. El enfoque se definió en función del dominio ya que el estudio se centra en las funciones de las NSOs dedicadas a la producción de estadísticas para desarrollo y los posibles tipos de modelos de madurez. Una vez definido el enfoque, se utilizaron dos técnicas para la identificación de actores: se realizó un análisis de stakeholders a partir de la bibliografía y se recolectaron opiniones de expertos. El resultado de estos ejercicios identificó actores del sector académico, privado, gubernamental y de la sociedad civil cuyas contribuciones son relevantes para la definición del modelo.

En la fase de *diseño* se tomaron decisiones que determinaron la arquitectura del modelo y sus componentes, los niveles de madurez y las dimensiones que se evaluarán. La definición de los niveles se realizó utilizando un enfoque vertical de abajo hacia arriba a partir del cual se identificaron los requerimientos y las medidas que informaron la definición de los niveles. Para la identificación de las dimensiones (actualmente en desarrollo) se está utilizando un proceso en cascada ya que el análisis iterativo de capas con distinto nivel de detalle facilita identificar el alcance y diferenciar las dimensiones en un dominio tan complejo. Para la definición del *contenido* se utilizará una combinación de métodos exploratorios como técnicas de Delphi, grupos nominales y entrevistas.

Las etapas de *validación*, *implementación* y *mantenimiento* dependerán en gran medida de las colaboraciones que se puedan lograr. Se intentarán establecer colaboraciones con NSOs de distintos países y con instituciones estadísticas internacionales. El proyecto ha despertado interés en el grupo de Datos para Desarrollo del Banco Mundial y se están explorando oportunidades con la Comisión Estadística de Naciones Unidas. En caso de que las colaboraciones esperadas no puedan concretarse, el desarrollo de estas etapas se basará en datos abiertos provistos por distintos gobiernos, como por ejemplo los del Reino Unido [16].

3. RESULTADOS OBTENIDOS/ESPERADOS

Los principales resultados obtenidos hasta el momento incluyen una revisión extensa de la bibliografía existente que describe las acciones y herramientas existentes para mejorar las estadísticas de indicadores sociales, y que respalda la hipótesis de que más investigación y soluciones son necesarias de manera urgente; la formulación de una versión preliminar de un CMM para determinar y mejorar la madurez de la capacidad de las organizaciones responsables de producir y reportar datos dentro de los ecosistemas nacionales; y un conjunto de recomendaciones para superar los obstáculos identificados dentro del dominio del control de indicadores sociales. Algunos de estos resultados fueron publicados en una conferencia internacional sobre tecnología, negocios y políticas para la administración y el análisis de datos organizada por la Unión Internacional de Telecomunicaciones [8]. Un artículo donde se define el problema de investigación y su alcance, las preguntas y los objetivos de investigación, la definición del marco teórico y el diseño de la investigación ha sido aceptado para publicación en una conferencia internacional sobre gobierno electrónico. Otro artículo que investiga las sinergias entre indicadores de gobierno digital y los indicadores de los SDGs ha sido presentado en una revista de gobierno electrónico.

Los próximos pasos incluyen una revisión en profundidad de las soluciones existentes para identificar las mejores prácticas para la producción de estadísticas para desarrollo y continuar con el desarrollo del CMM integrando los resultados de dicha revisión.

4. FORMACIÓN DE RECURSOS HUMANOS

Las tareas de investigación y desarrollo de este proyecto servirán de base para una tesis de Doctorado en Ciencias de la Computación a presentarse en la UNS, así como para tesinas de grado y el desarrollo de un curso de postgrado.

5. BIBLIOGRAFÍA

[1] “Last Mile Data Enablement and Collaboration, and Building Trust in Indicators Data | UNU

Computing and Society | UNU Computing and Society.” [Online]. Available: <http://cs.unu.edu/research/sdgs/>. [Accessed: 13-Mar-2018].

- [2] “Líneas de investigación – LISSI DCIC CIC.” [Online]. Available: https://lissi.cs.uns.edu.ar/?page_id=26. [Accessed: 18-Mar-2018].
- [3] “Departamento de Ciencias e Ingeniería de la Computación | UNS.” [Online]. Available: <https://cs.uns.edu.ar/home/>. [Accessed: 18-Mar-2018].
- [4] “UNSD - United Nations Statistical Commission.” [Online]. Available: <http://unstats.un.org/unsd/statcom>. [Accessed: 31-Jan-2017].
- [5] M. Thinyane, “Small Data for SDGs Community-Level Action and Indicators Monitoring.” 2016.
- [6] Independent Expert Advisory Group on a Data Revolution for Sustainable Development, “A World that Counts: Mobilising the Data Revolution for Sustainable Development,” 2014.
- [7] “Metrics & Indicators - Business for 2030.” [Online]. Available: <http://www.businessfor2030.org/metrics-indicators/>. [Accessed: 31-Jan-2017].
- [8] I. Marcovecchio, M. Thinyane, E. Estevez, and P. Fillostrani, “Capability Maturity Models Towards Improved Quality of the Sustainable Development Goals Indicators Data,” in *Proceedings of the 9th ITU Kaleidoscope academic conference: Challenges for a data-driven society, Nanjing, China, 27-29 November, 2017*, pp. 65–72.
- [9] K. Jugdev and J. Thomas, “Project Management Maturity Models: The Silver Bullets of Competitive Advantage?,” *Proj. Manag. J.*, vol. 33, no. 4, pp. 4–14, 2002.
- [10] T. De Bruin, R. Freeze, U. Kaulkarni, and M. Rosemann, “Understanding the Main Phases of Developing a Maturity Assessment Model,” in *Australasian Conference on Information Systems (ACIS)*, 2005, pp. 8–19.
- [11] R. Elio, J. Hoover, I. Nikolaidis, M. Salavatipour, L. Stewart, and K. Wong, “About Computing Science Research Methodology,” p. 9, 2011.
- [12] H. A. Simon, *The Sciences of the Artificial (3rd Ed.)*. Cambridge, MA, USA: MIT Press, 1996.
- [13] A. R. Hevner, “A Three Cycle View of Design Science Research,” *Scand. J. Inf. Syst.*, vol. 19, no. 2, pp. 87–92, 2007.
- [14] E. Ferrace, “Action Research,” *Rev. Gaucha Enferm.*, vol. 31, no. 3, pp. 567–574, 2010.
- [15] R. Wendler, “The maturity of maturity model research: A systematic mapping study,” *Inf. Softw. Technol.*, vol. 54, no. 12, pp. 1317–1339, 2012.
- [16] “data.gov.uk.” [Online]. Available: <https://data.gov.uk/>. [Accessed: 18-Aug-2017].

Los Procesos de Negocio en las Smart City: Un nuevo Paradigma

Carlos Salgado⁽⁺⁾; Mario Peralta⁽⁺⁾; Alberto Sánchez⁽⁺⁾; Saldarini Javier^(*), Carrizo Claudio^(*),
Armando Silvana^(*), Trasmontana Julio^(*)

⁽⁺⁾ Departamento de Informática - Facultad de Ciencias Físico-Matemáticas y
Naturales - Universidad Nacional de San Luis
Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina
e-mail: rivoira.ale@gmail.com, {csalgado, mperalta, alfanego}@unsl.edu.ar

^(*) Grupo de I+D Calidad de Software - Facultad Regional San Francisco
Universidad Tecnológica Nacional
Av. de la Universidad 501 - San Francisco - Córdoba - Tel. 03564-421147
{saldarinijavier, cjcarrizo77}@gmail.com

Resumen.

Desde la perspectiva de la complejidad de los procesos de negocio, una de las tecnologías más significativas para soportar su automatización son los Sistemas de Gestión Workflow. Para favorecer y flexibilizar dichos sistemas, es fundamental tener herramientas para medir su calidad. Debido a la globalización de la información y la comunicación, las empresas están migrando sus procesos a la nube y a la utilización de las nuevas tecnologías, como *IoT* y las redes sociales, lo que permite a las empresas tener una mejor disponibilidad de información.

La recolección de datos e información, provenientes de la nube y los dispositivos *IoT*, producen un cambio en la manera de apropiarse de dicha información. Lo que permite que las empresas y/u organizaciones pueden tener workflows ordenados y sincronizados, para un negocio de calidad.

En base a ello, se está trabajando en modelos ad-hoc para medir el grado de inteligencia de una ciudad. Que facilite la evaluación de procesos PN y sus workflow. Y beneficie su efecto en la nube, redes sociales y en los procesos de Marketing y Comunicación.

Esto presupone un nuevo paradigma de negocio que redundará en mejores oportunidades.

Palabras clave: Ciudades Inteligentes,

Smart City, Internet de las Cosas, *IoT*, Nube, Procesos de Negocio, Workflow, Modelos

Contexto

La presente línea de I+D se enmarca en el Proyecto de Investigación: Gestión de Proyectos de Software: *Los Modelos de Calidad como Soporte a los Procesos y Productos Software*. Esta propuesta está contextualizada en el trabajo colaborativo entre dos grupos de investigación: el Grupo de I+D Calidad de Software, perteneciente a la Facultad Regional San Francisco, Universidad Tecnológica Nacional y, por otra parte, los integrantes del Laboratorio de Calidad e Ingeniería de Software (LaCIS) – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis;

La evaluación, homologación y financiamiento del proyecto estuvo a cargo de la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional, el mismo es reconocido bajo el código: IAN4895.

Introducción

En este contexto, el enfoque de ciudades inteligentes es más que una opción o un planteamiento de futuro, es una necesidad en el presente para la mayor parte de los núcleos urbanos que, de una manera más o menos inminente, se enfrentan a problemas derivados del aumento constante de sus poblaciones, lo

que conlleva a, por ejemplo: (1) Problemas en la atención y calidad de los servicios públicos, que no pueden responder a la sobredemanda en aspectos especialmente sensibles como la sanidad, la educación, seguridad, servicios sociales y/o servicios en general; (2) Ralentización de la atención a la ciudadanía por parte de una administración saturada de peticiones y gestiones; (3) Problemas de movilidad derivados del aumento del tráfico y del uso de los vehículos particulares; (4) Saturación del mercado laboral y de los modelos clásicos de ocupación; entre otros

Los problemas mencionados anteriormente son los que persigue resolver este movimiento en donde, a través de las soluciones inteligentes y de las soluciones TIC como el Big Data [1] e Internet de las Cosas [2], promueven una nueva forma de entender las relaciones de los ciudadanos y su entorno urbano e impulsan, entre otros, los siguientes beneficios [3]: (1) Mejora de la calidad de los servicios públicos que se vuelven más eficientes; (2) Existencia de una comunicación real entre la ciudadanía y su ciudad; (3) Reducción de los costos de los servicios a través de soluciones Smart.

Esta transformación de una ciudad a una ciudad inteligente promueve una mejora de la relación ciudad-ciudadanos y convierte el entorno urbano en un lugar accesible, acogedor y sostenible, mejorando la calidad de vida de sus habitantes.

A la hora de poner las bases para la implantación de un modelo Smart City, el primer requisito por el que tienen que apostar las ciudades es el de la conectividad de todo el entorno urbano que, a partir de ahora debe entenderse en un concepto global y que va a contar con los siguientes elementos tecnológicos:

1. Administración y provisión de soluciones de acceso libre sobre plataformas abiertas para fomentar la

toma de decisiones compartidas con el ciudadano en el centro de la toma de decisiones.

2. Ciudadanos inteligentes que participen en los procesos de toma de decisiones y estén concienciados de los beneficios que les reporta el formar parte de una ciudad inteligente.
3. Sistemas de sensorización (Smart Sensors) aplicados en mobiliario público que reporten datos de forma constante para una gestión eficiente en el tratamiento de residuos, los sistemas de iluminación con paneles fotovoltaicos, la gestión eficiente del tráfico con señalización inteligente, en definitiva, Ciudades Conectadas.
4. Uso de soluciones Big Data para la recogida, interpretación y respuesta e interconexión constante con la ciudadanía a través de redes interconectadas (Smart Grid) o elementos del mobiliario público que permiten reportar cualquier incidencia, consulta o reclamo (Internet de las Cosas).
5. Edificios inteligentes (Smart Buildings) que apuesten por el ahorro energético y la generación de sus propias soluciones para su autoabastecimiento energético. Estas soluciones Smart City se trasladan a los hogares a través de sistemas de domótica que gestionan de forma eficiente nuestras viviendas.
6. Transporte sostenible, tanto en los sistemas de movilidad pública, como en los privados a través de vehículos eficientes (coche eléctrico, redes de alquiler de bicicletas, transporte público eficiente, etc.)
7. Gestión del urbanismo y ordenación de las ciudades basada en el aprovechamiento de los recursos y su gestión eficiente a través de soluciones como las grandes ciudades que repercuten en la mejora de la calidad de vida de los ciudadanos.

8. Elementos y herramientas de Ciberseguridad que garanticen la privacidad en las comunicaciones y la gestión de los datos personales que reportan los ciudadanos y que custodian la Administración.

Una ciudad inteligente se basa, principalmente en los siguientes subsistemas [3]:

- Generación distribuida: Consiste en que la ciudad inteligente posea generación eléctrica repartida por el territorio: el abastecimiento es individualizado (micro-generación), no centralizado.
- Smart Grids: Se conoce como Smart Grids a las redes inteligentes interconectadas, las cuales poseen una circulación bidireccional de datos entre el centro de control y el usuario.
- Smart Metering: Se trata de la medición inteligente de los datos de gasto energético de cada usuario, a través de telecontadores donde se realizan las lecturas a distancia y en tiempo real.
- Smart Buildings: Como modelo de eficiencia, los edificios deben ser inteligentes. Edificios domóticos que respetan el medio ambiente y que poseen sistemas de producción de energía integrados.
- Smart Sensors: Los sensores inteligentes tendrán la función de recopilar todos los datos necesarios para hacer de la ciudad una Smart City. Son parte fundamental para mantener la ciudad conectada e informada, y hacer que cada subsistema cumpla su función.
- eMobility: Implantación del vehículo eléctrico, y los respectivos puestos de recarga públicos y privados.
- Smart Citizen: Los ciudadanos son, sin duda, la parte fundamental de una Smart City, ya que sin su participación activa no es posible llevar a cabo estas iniciativas.
- Tecnologías de la información y la comunicación (TIC): En sí, para

nosotros, es la más importante ya que las tecnologías de la información son las que ayudarán a controlar los diferentes subsistemas que componen la Smart City, mediante las cuales los ciudadanos y las entidades administrativas pueden participar activamente en el control de la ciudad.

Surge, con lo antedicho, la necesidad de saber cuándo una ciudad es inteligente, o el grado de inteligencia de la misma. Para determinar si una ciudad es o no inteligente, generalmente, se toma como punto de partida un modelo en torno a seis características: *Economía, Gestión de gobierno, Ciudadanía, Entorno, Calidad de vida y Movilidad*. A partir de este modelo, se han desarrollado objetivos agrupados en una serie de servicios para los que se han descrito las principales tecnologías que disponemos en la actualidad. De esta forma, se obtuvo una serie de plataformas de servicios, que se deben integrar y conectar entre sí, facilitando que el ciudadano forme parte activa del proceso de gestión de su ciudad [4]. En otras palabras, se busca modernizar la gestión de las ciudades, fomentando una mayor interacción entre las instituciones, los ciudadanos y los sistemas informáticos [5].

Desde otro punto de vista, El trabajar en la nube, y más aún con IoT, lleva a tener que diseñar estrategias de marketing en internet. Hoy en día el mundo digital no deja de ofrecer nuevas herramientas y soluciones para incrementar la efectividad de muchas de las técnicas usadas para la captación de nuevos clientes. Una de estas mejoras en los últimos años ha sido la posibilidad de automatizar muchas acciones de marketing para acortar tiempos de espera y, así, aumentar la satisfacción de los usuarios que interactúan con una página web. Una herramienta que permite esta automatización es el uso de las diversas redes sociales existentes en la actualidad. Estas redes ayudan en gran medida a

fidelizar a los clientes.

Por su parte, desde Gartner [6], donde investigan el mercado y asesoran a sus clientes, observan otro cambio disruptivo y afirman que el IoT creará "nuevas dinámicas de marketing, ventas y servicio al cliente". En un gran número de transacciones, las interacciones serán directamente entre cosas conectadas, sin considerar decisiones personales. Un ejemplo es el automóvil, que puede coordinar directamente los servicios de mantenimiento, sin involucrar al propietario en la mayor parte del proceso. Otro ejemplo es encontrar estacionamiento [Tesis de Imperiale] o mejor aún, encontrar uno en el que se pueda recargar un auto eléctrico. El automóvil podrá negociar directamente esta acción.

Desde otro punto de vista, la tendencia actual va de la mano de los dispositivos móviles y la ubicuidad, y abre la posibilidad de numerosas nuevas alternativas", y de brindar soluciones a la medida y la personalización del consumidor.

Ante la globalización de las comunicaciones, la información y el comercio electrónico, y la necesidad de las organizaciones de mantener su negocio altamente competitivo, las empresas están migrando sus procesos de negocio a la nube [7]. Esto se debe a que estar en la nube significa movilidad, seguridad, escalabilidad y elasticidad. Es decir, poder dimensionar los servicios a lo que se precisa, incluso programando necesidades periódicas, despreocuparse de los equipos y las tecnologías, del mantenimientos y reparaciones, de la compra de equipos, todo lo cual es atendido por el proveedor del servicio.

No obstante las ventajas que proveen la nube y las nuevas tecnologías, no siempre es posible, o no siempre las organizaciones están dispuestas a hacerlo, subir sus procesos a la nube y adaptarse a los nuevos paradigmas de comunicación, ya que no confían en la

seguridad de la misma o porque consideran que no les será rentable. Por ello, los dueños de los procesos necesitan tener un medio que les permita evaluar la conveniencia de subir todos o parte de sus procesos a la nube y, en caso de subir alguno de ellos, decidir cuáles. Al igual que decidir si incorporar el uso de las redes sociales y las nuevas tecnologías que pueden surgir para mejorar sus procesos de comunicación y marketing [7].

Otro de los aspectos fundamentales en todo proceso workflow, es tener medios que permitan medir sus elementos más relevantes con el fin de detectar las áreas y aspectos a mejorar del proceso y, de esta manera, promover su mejora continua. Para ello, creemos que es necesario proveer un marco que permita realizar dicha medición. De acuerdo a esta necesidad, siguiendo la metodología propuesta en [8], hemos definido un conjunto de métricas elementales que sirven como indicadores de la complejidad estructural de los modelos de procesos workflow [9].

Desde otra perspectiva, las empresas están dirigiendo su mirada a la administración de sus procesos en la nube e IoT. Desde el punto de vista de la administración de los procesos, un workflow puede ser controlado de forma manual, informatizada, o como una combinación de ambos métodos. Un esquema workflow permite a las empresas organizar las tareas y recursos a través de reglas que facilitan el control de los PN de la empresa. Es decir que, con un proceso workflow se logra un control total y absoluto de todas las tareas. En este sentido, las empresas interesadas en llevar su negocio a la nube, necesitan mantener ese control. Por ello, es fundamental que los procesos workflow sean de alta calidad [10].

Cada vez más empresas, de distintas industrias, ven revolucionado su negocio gracias a las nuevas tecnologías. La combinación de IoT, la nube y las

ciudades inteligentes, plantea la necesidad de tener herramientas, métodos, modelos, etc., para estudiar los PN de las distintas organizaciones, y que permitan hacer evaluaciones y recomendaciones que ayuden en la toma de decisiones a los gobiernos para mejorar las condiciones de vida de las sociedades. Esto plantea un desafío en las necesidades de las distintas sociedades para lograr mejoras en la calidad de vida.

Resultados Obtenidos y Objetivos

Entre los resultado que se han obtenido en esta línea de investigación, se pueden detallar el modelo de calidad con base en [4, 5].

Se ha construido un modelo ad-hoc de calidad teniendo en cuenta las redes inteligentes como los aparatos inteligentes dentro de ellas. Teniendo características como Conectividad y geo localización; Tráfico Vehicular; Redes Eléctricas, Reducción del consumo de energía, etc.

Se están definiendo métricas e indicadores y adecuando a esta nueva realidad de las ciudades inteligentes, internet de las cosas, la nube y dentro de este contexto están los procesos de negocio que dan impulso a las organizaciones. Se está trabajando en la definición de distintos indicadores que nos permitan evaluar la calidad de las ciudades inteligentes con sus aparatos inteligentes y los modelos de los flujos de trabajos que representan estos nuevos paradigmas de negocio y marketing.

Formación de Recursos Humanos

Bajo esta línea de investigación, en el grupo se están desarrollando una tesis de maestría en Calidad de Software, como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Ciencias de la

Computación e Ingeniería en Informática.

Referencias

- [1] M. Bouskela, M. Casseb, S. Bassi, C. De Luca, and M. Facchino, "La ruta hacia las smart cities: Migrando de una gestión tradicional a la ciudad inteligente," *Monografía del BID*; 454
- [2] C. Doukas, "Building Internet of Things with the Arduino," *US Trade Paper*, 2012.
- [3] V. Vitolins, "Business Process Measures," in *Int. Conference on BALTIC DB&IS. Riga, Latvia.*, 2004, pp. 186-197.
- [4] S. Colado, A. Gutiérrez, C. J. Vives, and E. Valencia, "SMART CITY: Hacia la gestión inteligente.," *S.A. MARCOMBO*.
- [5] F. P. Digital, "País Digital: Smart Cities." <http://www.paisdigital.org/PD/smart-cities/>
- [6] Gartner. <https://www.gartner.com/en>.
- [7] M. Peralta, C. Salgado, L. Baigorria, G. Montejano, and D. Riesco, "Procesos Workflow en la Nube: Una Propuesta para Evaluar su Migración.," in *Paper presented at the 3er. CoNaISSI 2015*.
- [8] M. Serrano, M. Piattini, C. Calero, M. Genero, and D. Miranda, "Un método para la definición de métricas de software.," in *1er Workshop en Métodos de Investigación y Fundamentos filosóficos en Ingeniería del Software y Sistemas de Información (MIFISIS'2002)*, 2002, pp. 65-74.
- [9] N. Debnath, M. Peralta, C. Salgado, L. Baigorria, G. Montejano, and D. Riesco, "Metrics and Performance Indicators to Evaluate Workflow Processes on the Cloud," in *12th AICCSA 2015*, 2015.
- [10] N. Debnath, M. Peralta, C. Salgado, G. Montejano, D. Riesco, and M. Berón, "Metrics for Evaluation of Workflow Models: An Experiment for Validation," *Journal of Computational Methods in Science and Engineering.*, vol. 12 2012.

Un Framework para la Evaluación de Modelos Conceptuales de Procesos de Negocio

Carlos Salgado, Mario Peralta, Daniel Riesco, Lorena Baigorria, Germán Montejano,

Departamento de Informática Facultad de Ciencias Físico-Matemáticas y Naturales Universidad Nacional de San Luis

Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina

e-mail: {csalgado, mperalta, driesco, flbaigor, gmonte}@unsl.edu.ar

Resumen

El objetivo de un framework para la evaluación de modelos es brindar a las organizaciones un medio que les ayude a mantener información objetiva y precisa acerca de la mantenibilidad, entendibilidad, acoplamiento y cohesión de los modelos, facilitando la evolución de los Procesos de Negocio de las empresas involucradas en una mejora continua. Además, proporciona soporte a la gestión de PNs al facilitar la evaluación temprana de ciertas propiedades de calidad de sus modelos. Con ello, las organizaciones se ven beneficiadas garantizando el entendimiento y la difusión de los PN y su evolución, y reduciendo el esfuerzo necesario para cambiar los modelos con la consecuente reducción de los esfuerzos de mantenimiento y mejora.

En esta línea de investigación, se propone un framework para la evaluación de modelos conceptuales de PN centrado en la aplicación de uno de dos métodos de evaluación [1]. Ambos métodos permiten enfrentar el mismo problema desde dos enfoques diferentes: uno referido a lo numérico y el otro que se acerca a expresiones lingüísticas similares al lenguaje cotidiano. Ambos aportan resultados importantes a distintas áreas, dependiendo de la perspectiva a analizar, dotando al framework de un valor agregado para el análisis de los modelos conceptuales de PN.

Palabras clave: Frameworks de modelado, Procesos de Negocio, Modelado de Procesos de Negocio, Modelos Conceptuales.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

Introducción

El modelado del Proceso de Negocio constituye una etapa fundamental para lograr el objetivo de un Proceso de Negocio (PN), es decir, como lo establece la definición de PN, obtener resultados beneficiosos para los clientes u otros interesados [2]. Esto se debe a que un modelo de PN describe cómo funciona el negocio [3], es decir, describe las actividades involucradas en el negocio y cómo se relacionan e interactúan con los recursos necesarios para lograr los objetivos del proceso. Desde este punto de vista, el modelo de PN se utiliza para capturar, documentar o rediseñar PNs [4].

Estos modelos, presentan una visión global de la organización que permite

entender mejor la dinámica de la empresa y las relaciones que se dan en su interior y con su entorno. Esto se da tanto en el ámbito que refiere a los clientes, como a sus proveedores y/o prestadores de servicios. El modelado del negocio es la técnica por excelencia para alinear los desarrollos con las metas y objetivos de las organizaciones, puesto que los modelos cumplen un rol fundamental en la especificación de los procesos de negocio.

En la literatura, [5, 6, 7] por mencionar algunos, se pueden encontrar diversas conceptualizaciones de los PN. No obstante, teniendo presente las diversas definiciones de PN existentes, se puede decir que, normalmente, un PN: (i) Está asociado con objetivos operacionales y relaciones de negocio, (ii) puede estar contenido completamente dentro de una unidad organizacional o puede abarcar diferentes organizaciones, (iii) tiene condiciones definidas que disparan su inicio, (iv) produce salidas definidas en su finalización, (v) puede involucrar interacciones formales o relativamente informales entre los participantes, y (vi) puede consistir de actividades manuales y/o automatizadas.

El desarrollo de modelos conceptuales constituye una parte del esfuerzo de implementación de un PN. Sin embargo, es una de las tareas claves en las primeras etapas del ciclo de vida de los PNs. Son utilizados, principalmente, como herramientas o medios para que, los distintos tipos de participantes, puedan entender fácilmente el proceso que dichos modelos representan. Además, son empleados como punto de partida a la hora de realizar cambios y adaptaciones de los procesos a las nuevas necesidades de la empresa. Por ello, la calidad de los mismos es de vital importancia para que ayuden a mejorar el desempeño y evolución de la organización y no se conviertan en un factor de riesgo. Bajo estas consideraciones, se propone un framework para evaluar modelos conceptuales de PN. El objetivo es proveer a las organizaciones un medio que les ayude a estudiar la calidad de sus Modelos de PN desde el punto de vista de su **entendibilidad** y su **adaptabilidad** a los cambios.

Respecto de dichas características, la **entendibilidad** permite al usuario

entender/comprender si los MPN son adecuados y cómo usarlos en tareas y condiciones de uso particulares. Es decir, brinda un indicativo acerca de cuán fácil es aprender a leer e interpretar dichos modelos para poder comprender la realidad que ellos están representando.

La **adaptabilidad**, por su parte, representa la capacidad del modelo para ser modificado efectiva y eficientemente, debido a necesidades evolutivas, correctivas o perfectivas. Es decir, que sea modificado de forma efectiva y eficiente sin introducir defectos o degradar su entendimiento y comprensión.

En función de lo expresado, se propone un framework destinado al análisis y estudio de los Modelos de PN desde la perspectiva de las características de calidad esperadas de un Modelo de PN y el cual se centra en el uso de la lógica continua o la lógica difusa, dependiendo de las características de los modelos estudiados y de los procesos que dichos modelos representan.

Líneas de Investigación, Desarrollo e Innovación

En base a lo expresado previamente, en esta línea de investigación se trabaja en el estudio, definición y uso de un framework para la evaluación de modelos conceptuales de procesos de negocio el cual se centra en la aplicación de distintos métodos de evaluación para diversos sistemas del mundo real. En particular, se proponen dos métodos para evaluar modelos conceptuales de PN: (i) Un método basado en la lógica continua [8]; y (ii) un método basado en la lógica difusa [9]. La selección y aplicación de uno u otro método, dependerá de las características de los modelos estudiados y de los procesos que dichos modelos representan.

Resultados Obtenidos y Objetivos

El framework propuesto, es un medio/herramienta que propone el estudio y análisis de los modelos de PN

de una institución y/u organización. El objetivo principal es guiar el desarrollo de dichos modelos mediante la especificación de requisitos y evaluación de características de calidad. Este framework está conformado por tres partes, cada una de las cuales complementa o colabora con la anterior. Estas partes se resumen como:

1. Aplicar el Analizador sintáctico al modelo, para determinar la corrección sintáctica del mismo.
2. Seleccionar y aplicar el método de evaluación, según la realidad o las necesidades a evaluar:
 - a. Método basado en operadores de lógica continua [10], o
 - b. Método basado en lógica difusa [11].
3. Analizar Resultados y generar reportes y recomendaciones.

El Analizador Sintáctico chequea la correcta conformación del modelo motivo de estudio. En esta etapa se corrobora, por ejemplo, si el modelo cumple o satisface las buenas prácticas o guías de modelado para modelos conceptuales de PN [12].

El cumplimiento de las guías de modelado por parte de los modelos de PN redundan en que los modelos sean entendibles y adaptables a las necesidades de las organizaciones. Facilitando la tarea de los distintos actores que intervienen en el proceso de modelado conceptual de los PN.

Una vez que los modelos sean correctos, se procede a hacer el estudio y/o análisis a través del uso de una (o ambas) de las dos alternativas que nos provee el framework. Es decir, optar por trabajar con: (i) *operadores de la lógica continua*, o (ii) *acercarnos más al lenguaje natural y propio de los seres humanos a través del uso de la lógica difusa, o algunas de sus alternativas*, o, en caso de no ser claro cuál de las alternativas es más adecuada para la situación particular, se pueden aplicar

ambas alternativas y luego realizar un análisis comparativo de los resultados obtenidos.

Uno de los aspectos primordiales de todo método de evaluación es mostrar que el mismo es de utilidad práctica. Para alcanzar este objetivo, y siguiendo la clasificación propuesta en [13], se aplicó el framework a un *caso de estudio* para analizar los modelos de PN de una empresa del medio. La decisión de utilizar casos de estudio para la validación del framework se debió a que, al evaluar un método para la evaluación de modelos de PN, en general, no se tiene un control absoluto de las variables a evaluar. Esto se debe a que, en la mayoría de los casos, dichas variables dependen de la realidad particular que se está estudiando. Por este motivo, se consideró más adecuada la aplicación de casos de estudio que la realización de experimentos, en los que es necesario tener un mayor control de las variables intervinientes, o el desarrollo de encuestas, para las cuales se debería contar con un cierto historial de aplicación del método y la opinión de quienes lo utilizaron.

En función de lo expresado, el caso de estudio mencionado, fue la aplicación del framework para la evaluación de los modelos de PN de una empresa del medio, la cual pretende posicionarse satisfactoriamente con competitividad en el mercado.

Cabe destacar que, en este caso, se optó por utilizar el método basado en la lógica difusa, ya que éste permite acercarse más a la manera de pensar de los seres humanos, permitiendo representar las ambigüedades que surgen en cuanto a la interpretación de las distintas reglas de negocio que pueden surgir.

De la aplicación del método, en el análisis de los distintos modelos de la empresa motivo de estudio, se detectó que el modelo propuesto, en general, era medianamente comprensible, es decir, no se adaptaba totalmente a los estándares de calidad de los modelos

conceptuales. Principalmente porque no era comprensibles en su totalidad.

A partir de este análisis y con el resultado obtenido, se procedió a la organización y acomodamiento de las componentes del modelo en un intento de hacerlos más entendibles para los distintos actores de la organización. Luego de reacondicionarlos se presentaron ante la gerencia de la empresa. La misma, tomó los nuevos modelos y se los cotejó contra los procesos reales de la organización. De dicha observación, se detectó que los procesos de negocio se adecuaban a los nuevos modelos, los cuales fueron desarrollados en función de las especificaciones de requerimientos realizados por la gerencia, sin tener contacto previo con el modelo del proceso con que ya contaba la empresa. Este trabajo llevó a la gerencia a detectar que, si bien el proceso se adaptaba a su modelo, el modelo no se adecuaba a la realidad de sus requerimientos de negocio. Por lo tanto, se debía hacer una reestructuración en la puesta en ejecución del proceso bajo estudio de la empresa.

Formación de Recursos Humanos

Bajo esta línea, en el grupo de investigación se presentó un trabajo final de Especialización en Ingeniería de Software, [14], y una tesis de Maestría en ingeniería de Software referente al Modelado de Procesos de Negocio [15]. Actualmente, se está trabajando en una tesis doctoral. Además, se está desarrollando una tesis de maestría en lo referente, como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Ciencias de la Computación y para la Ingeniería en Informática.

Referencias

[1] N. Debnath, C. Salgado, M. Peralta, D. Riesco, L. Baigorria, G. Montejano, I. Lee, and H. Lee,

"Quality of Business Processes Conceptual Models: Fuzzy Logic vs. Continuous Logic Operators," in *30th International Conference on Computer Applications in Industry and Engineering (CAINE 2017)*, San Diego, California, USA, 2017, pp. 205-210.

- [2] A. Sharp and P. McDermott, "Workflow Modeling: Tools for Process Improvement and Application Development.," *London: Artech House*, 2001.
- [3] T. Dufresne and J. Martin, "Process Modeling for e-Business," *Spring 2003, INFS 770 - Methods for Informations Systems Engineering: Knowledge Management and E-Business. George Mason University*, 2003.
- [4] B. Mora, F. Ruiz, F. García, and M. Piattini, "Experiencia en transformación de modelos de procesos de negocios desde BPMN a XPDL.," *IDEAS*, 2007.
- [5] C. Jiménez, L. Farías, and F. Pinto, "Análisis de Modelos de Procesos de Negocios en relación a la dimensión informática.," *Revista Electrónica del DIICC. <http://www.inf.udec.cl/revista/edicones/edicion9/cjimenez.pdf>*, 2004.
- [6] G. Sparks, "An Introduction to UML.," *The Business Process Model. Sparx Systems, www.sparxsystems.com.au*, 2000.
- [7] Z. Irani, V. Hlupic, and G. M. Giaglis, "Business Process Re-Engineering: a Modeling Perspective," *International Journal of Flexible Manufacturing Systems*, pp. 99-104., 2001.
- [8] J. J. Dujmovic, "Quantitative Methods for Software Evaluation," *Lecture Notes, Graduate Software Engineering Program. Universidad Nacional de San Luis, San Luis, Argentina.*, 1998.
- [9] L. A. Zadeh, "Fuzzy Sets," *Information and Contgrol*, vol. 8,

- pp. 338-353, 1965.
- [10] N. Debnath, C. Salgado, M. Peralta, M. Berón, D. Riesco, and G. Montejano, "MEBPCM: A Method for Evaluating Business Process Conceptual Models. A Study Case.," presented at the Ninth International Conference on Information Technology: New Generations (ITNG), Las Vegas, Nevada, USA, 2012.
- [11] N. Debnath, C. Salgado, M. Peralta, D. Riesco, L. Baigorria, and G. Montejano, "A Fuzzy Logic-based Method to Evaluate the Quality of Business Process Models," in *CATA 2017*, 2017.
- [12] J. Mendling, H. Reijers, and W. van der Aalst, "Seven Process Modeling Guidelines," *Information & Software Technology*, vol. 52(2), pp. 127-136, 2010.
- [13] C. Robson, *Real world research: A resource for social scientists and practitioners-researchers*: Blacwell, 1993.
- [14] C. Salgado, "Tesis de Especialización en Ingeniería de Software: El Modelado de Procesos de Negocio: Aplicando LSP para la Evaluación de Lenguajes de Modelado de Procesos de Negocio," Facultad de Ciencias Físico-Matemáticas y Naturales. UNSL, 2010.
- [15] C. Salgado, "Tesis de Maestría en Ingeniería de Software: MEMPN: Método para la Evaluación de Modelos Conceptuales de Procesos de Negocio ", Facultad de Ciencias Físico-Matemáticas y Naturales., Universidad Nacional de San Luis, San Luis, Argentina., 2013.

Implementación de una arquitectura de procesos como resultado de la aplicación del ciclo de vida BPM durante sus fases de configuración y ejecución

Marisa Pérez*, Juan Pablo Ferreyra*, Claudia Verino*, Diego Cocconi*

Departamento de Ingeniería en Sistemas de Información / Facultad Regional San Francisco / Universidad Tecnológica Nacional (UTN)
Av. de la Universidad 501, 2400, San Francisco, Córdoba, Argentina, (03564) 431019 / 435402
*{mperez, jpferreyra, cverino, dcoconi}@sanfrancisco.utn.edu.ar

Resumen

La gestión de procesos de negocio (BPM) está ganando mucha popularidad hoy en día. La misma se fundamenta en la representación explícita de los procesos de negocio en las organizaciones para posibilitar la aplicación de un ciclo de mejora continua en el que intervienen las fases de diseño y análisis, configuración, ejecución y evaluación (ciclo de vida de BPM). El propósito de este trabajo, que continúa lo expuesto en un trabajo anterior, es remarcar la importancia de las fases de configuración y ejecución de dicho ciclo para poner en valor la información que se genera mediante las trazas de los procesos que permiten medir performance mediante indicadores (KPI), para obtener realimentación que luego será utilizada en la mejora de los procesos.

Por lo expuesto, el principal objetivo de la presente línea de investigación en este momento es la implementación de la arquitectura de procesos definida en la fase de diseño y análisis, durante una etapa anterior del proyecto. Este objetivo implica configurar un motor de procesos para dar soporte a la ejecución de los modelos de procesos definidos en la arquitectura, configurar los componentes que sean necesarios para permitir el monitoreo de variables que posibiliten calcular los KPI asociados a la arquitectura de procesos y realizar la puesta en ejecución del sistema para que los usuarios finales puedan interactuar.

Una vez alcanzado dicho objetivo, se habrá obtenido una base de datos con todas las variables involucradas en la ejecución de las instancias de los procesos y los valores de los KPI reales, los cuales podrán ser contrastados con los esperados (determinados también en la fase de diseño y análisis). La base de datos también será fuente de los logs para realizar minería de procesos en un proyecto futuro.

Empleando estas dos fuentes de realimentación obtenidas durante el monitoreo y la minería de procesos, se obtiene información para validar y/o proponer una nueva arquitectura de procesos mejorada que resolverá las falencias identificadas en la actual. Esta nueva arquitectura será propuesta como mejora dentro del proyecto de investigación del cual surgió este estudio, completando la ejecución de una iteración completa del ciclo de vida BPM.

Palabras clave: gestión de procesos de negocio, mejora continua, arquitectura de procesos, ciclo de vida de BPM, sistemas de información.

Contexto

El presente trabajo se encuadra dentro del proyecto de investigación I+D UTN 4090 “Optimización organizacional basada en un modelo de gestión por procesos en la Secretaría Extensión y Cultura de la Universidad Tecnológica Nacional, Facultad

Regional San Francisco”. El mismo se encuentra homologado como proyecto de investigación y desarrollo por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional.

En el marco de dicho proyecto se propone promover BPM para la mejora del desempeño de las organizaciones, mediante la gestión por procesos. Para fomentar BPM en la institución, con el fin de mejorar su eficiencia y rendimiento, y lograr la alineación de los procesos de negocio con metas y estrategias, se ha continuado con su aplicación en la Secretaría de Extensión y Cultura de la UTN Facultad Regional San Francisco.

1. Introducción

La *gestión de procesos de negocio* (BPM, del inglés *Business Process Management*) está ganando mucha popularidad hoy en día. BPM permite a las organizaciones administrar y mejorar sus procesos de negocio, de manera que los mismos evolucionen y conduzcan a una optimización organizacional [1,2]. BPM no sólo se trata de mejorar la forma en que se llevan a cabo actividades individuales, sino de la gestión de cadenas enteras de eventos, actividades y decisiones que agregan valor a la organización y sus clientes. Estas cadenas de eventos, actividades y decisiones se denominan *procesos de negocio* [3]. Un proceso de negocio consiste en un conjunto de actividades que se llevan a cabo de manera coordinada en un entorno organizacional. Cada proceso de negocio contribuye a alcanzar uno o más objetivos de la organización [4].

BPM permite aplicar un ciclo de mejora continua que se conoce como *ciclo de vida de BPM* (Fig. 1), en el que intervienen las fases de *diseño y análisis*, *configuración*, *ejecución* y *evaluación*. En la fase de diseño y análisis se identifican los procesos de negocio de la organización, lo cual implica un análisis del estado actual de los procesos con el objetivo de detectar problemas existentes e identificar oportunidades de mejora para el (re)diseño de los mismos. Durante la fase de configuración,

se especifican aspectos necesarios y se implementan para que los modelos de procesos puedan ser interpretados por un *sistema de gestión de procesos de negocio* (BPMS, del inglés *Business Process Management System*). En la fase de ejecución, el BPMS permite la ejecución de los procesos configurados mediante un *motor de procesos*. Finalmente, en la fase de evaluación se analiza el resultado de la ejecución para identificar problemas y aspectos que puedan ser mejorados [5].



Fig. 1: Ciclo de vida de BPM [4].

En este trabajo se propone recalcar la importancia de las fases de configuración y ejecución (especialmente del monitoreo, dentro de esta última), como predecesoras de la fase de diseño y análisis, durante la cual se obtuvo una *arquitectura de procesos*. Una arquitectura de procesos (o *mapa de procesos*) define “*qué*” hace y hará en el futuro la organización. La arquitectura de procesos es una colección de procesos de negocio y sus interdependencias entre ellos [6]; estos procesos deben estar alineados a los objetivos de la organización.

Dicha arquitectura de procesos puede ser optimizada sobre la base de los resultados obtenidos en las etapas involucradas a lo largo de varias iteraciones (para el caso de estudio fueron propuestas dos iteraciones).

A modo de caso de estudio, se continuará ampliando el que corresponde al proyecto I+D homologado, el cual describe el trabajo en una secretaría de la Facultad Regional San

Francisco: la Secretaría de Extensión y Cultura. Una vez probada la efectividad de la propuesta, se propone y sugiere aplicarla a otras áreas de la Facultad para optimizarlas y lograr la visión global de la institución completa.

2. Líneas de Investigación, Desarrollo e Innovación

Este trabajo se encuadra dentro de una línea de investigación que abarca las siguientes áreas temáticas:

- Gestión de procesos de negocio.
- Mejora de procesos.
- Metodologías de análisis y diseño de procesos.

Particularmente, en esta etapa de investigación y desarrollo, se hace foco en la fases de configuración, ejecución y evaluación (monitoreo) de procesos.

3. Objetivos y Resultados Esperados

El principal objetivo durante esta etapa del proyecto es, a partir de la arquitectura de procesos definida previamente mediante el enfoque descrito en un trabajo previo [7], realizar la implementación de la misma, logrando la ejecución de los procesos, el cálculo de los KPI, el monitoreo de los procesos y la generación de los *logs* necesarios que servirán como entrada a la fase de evaluación del ciclo de vida de BPM.

Este objetivo implica configurar un *motor de procesos* para dar soporte a la ejecución de los procesos definidos en la arquitectura, generando las instancias de proceso correspondientes. El motor de procesos es parte de un BPMS y proporciona diferentes funcionalidades: (1) crear instancias de proceso (también llamadas *casos*) a partir de modelos ejecutables; (2) distribuir el trabajo de los participantes (usuarios) del proceso en orden de ejecutarlo; (3) recolectar y almacenar datos requeridos para la ejecución

del proceso y delegar las actividades (automáticas) a los distintos sistemas de software existentes en la organización; (4) monitorear el progreso de las diferentes instancias [4]. Se evaluaron diferentes alternativas (*jBPM*¹, *Bonita*² y *Bizagi*³) y se ha optado por el motor de procesos de *Bonita* (versión de código abierto), debido a la flexibilidad brindada en cuanto a las API (interfaz de programación de aplicaciones, del inglés *Application Programming Interface*) y conectores para aplicaciones externas, y también a la amplia cantidad de documentación existente a la hora de realizar estas conexiones, que suelen involucrar diferentes sistemas de información.

Los modelos de procesos de la arquitectura deben adaptarse para convertirlos en modelos ejecutables, pues los modelos generados en la fase de diseño y análisis tienen un nivel de abstracción algo más elevado y carecen de detalles de implementación. Una vez configurados los modelos ejecutables, el motor de procesos puede crear las instancias necesarias de los procesos, interactuando con los usuarios y sistemas software. A medida que se ejecutan los procesos, deben determinarse los KPI asociados a los mismos. Los KPI (del inglés *Key Performance Indicators*) son indicadores de alto nivel mediante los cuales los ejecutivos intentan monitorear y asegurar que los objetivos, estrategias e iniciativas relacionadas son obtenidas [8]. A tal fin se construye una base de datos donde se colecta información sobre las instancias de proceso (y las instancias de cada tarea) desde el motor de procesos. Mediante los conectores apropiados, el motor de procesos se vincula a la base de datos y se van agregando (o actualizando) los registros a medida que se van ejecutando las diferentes instancias de los procesos y las tareas. Esta base de datos dará soporte al monitoreo de los procesos.

¹ www.jbpm.org

² www.bonitasoft.com

³ www.bizagi.com

Durante la etapa de ejecución (actualmente en curso), los usuarios finales interactuarán con el sistema y poblarán la base de datos de monitoreo. Una vez alcanzado este objetivo, se obtendrá una base de datos con toda la información de la ejecución de las instancias de los procesos. Los valores en esta base de datos servirán de variables de entrada para determinar los KPI reales, así podrán ser contrastados con los esperados, determinados en la fase de diseño.

Los KPI generalmente se agrupan en tres categorías: *costos del producto o servicio*, *eficiencia y calidad de las salidas* [8]. Para determinar los KPI relativos a costo, se usan tablas de recursos utilizados o involucrados en cada proceso definido en la arquitectura (pues dichos recursos serán propios de cada proceso), a partir de los cuales puede determinarse un costo. Con los KPI basados en tiempo es importante considerar las estampas de tiempo de inicio y fin tanto de tareas como de procesos, valores que se almacenan en las tablas “*actividad*” y “*proceso*” del esquema de la base de datos (Fig. 2), respectivamente. Finalmente, para los KPI de calidad, la información de todas las tablas es pertinente.

Respecto del monitoreo propiamente dicho, inicialmente se propuso el desarrollo de un sistema de monitoreo que tome los datos de la base de datos y componga los KPI, mostrándolos adecuadamente. Debido a la falta de flexibilidad que ello implica, en esta etapa del proyecto se está analizando un sistema de *Business Activity Monitoring* (BAM), el cual permitiría conectar la base de datos, definir indicadores y mostrarlos en un tablero. Actualmente se está trabajando con el producto *opensource Stream Processor* de WSO2⁴.

La base de datos será fuente también de los *logs* que se utilizarán para desarrollar el próximo proyecto I+D ya homologado, sobre minería de procesos, en la fase de evaluación del ciclo de vida de BPM.

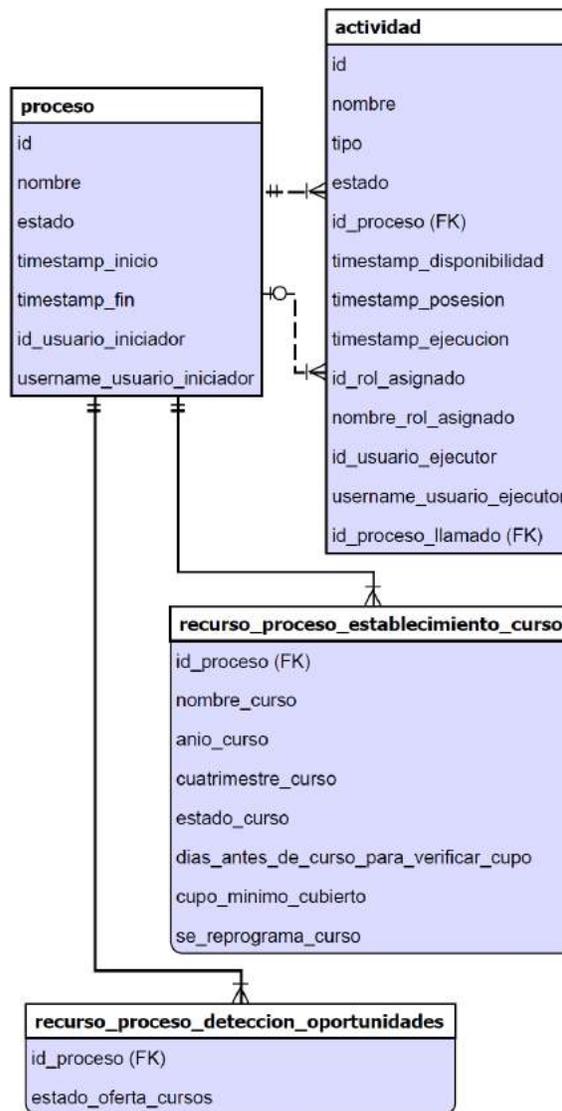


Fig. 2: Esquema parcial de la base de datos de monitoreo.

Debido a la naturaleza iterativa que conlleva implícitamente la mejora continua de los procesos, necesaria para lograr la plena optimización organizacional, se pretenden estructurar los resultados de la última fase, de forma que sirvan como entrada para una nueva iteración. En el trabajo previo que sirve de base para el presente [7], fueron propuestas dos iteraciones.

4. Formación de Recursos Humanos

El grupo de esta línea de investigación está conformado por docentes y alumnos de la carrera de Ingeniería en Sistemas de

⁴ <https://wso2.com/analytics>

Información. Entre los docentes, tres de ellos se encuentran en la etapa de desarrollo de sus tesis de maestría (dos en Ingeniería en Sistemas de Información y uno en Ingeniería de Software); otro está comenzando su tesis de doctorado (mención Ingeniería en Sistemas de Información). Todos ellos en temas altamente vinculados al área del proyecto. Además un integrante ya ha presentado su trabajo final de Especialización en Ingeniería en Sistemas de Información en un área referida al tema.

Como iniciativa del grupo, se prevé la capacitación y formación de recursos humanos, que contempla las siguientes actividades:

- Participación en cursos de actualización y posgrado en el área de estudio (procesos de negocio).
- Intercambio de ideas y conocimientos con personal de otras Facultades que investiguen en el área (Santa Fe, Mendoza, Rosario).
- Transferencia de tecnologías de procesos a otras áreas de la Facultad y a la industria local por medio de talleres, cursos y charlas.
- Dirección y asesoramiento sobre el área a interesados de la industria local.
- Incorporar a las cátedras del tronco integrador de la carrera Ingeniería en Sistemas de Información (Sistemas y Organizaciones, Análisis de Sistemas, Diseño de Sistemas y Administración de Recursos) contenidos sobre la gestión de procesos de negocio utilizando la experiencia y los conocimientos obtenidos en el presente proyecto.
- Involucrar alumnos de la carrera Ingeniería en Sistemas de Información en la realización de actividades del presente proyecto e incentivarlos para acercarse a propuestas de becas.

5. Referencias

[1] Duipmans, E. F., Pires, L. F. y da Silva Santos, L. O. B. (2014). "A transformation-based approach to business process management in the cloud". *Journal of grid computing*, 12(2), 191-219.

[2] Ferreyra, J. P., Roa, J., Cocconi, D., Perez, M., Verino, C., Villarreal, P. D. (2017). "Estado actual de la Gestión de Procesos de Negocio basada en Computación en la Nube". En: *5to. Congreso Nacional de Ingeniería Informática / Sistemas de Información (CONAISI)*, 2017, Santa Fe, Argentina.

[3] Dumas, M., La Rosa, M., Mendling, J. y Reijers, H. A. (2013). "*Fundamentals of business process management*" (Vol. 1, p. 2). Heidelberg: Springer.

[4] Weske, M. (2012). "*Business process management: concepts, languages, architectures*", 2da. ed. Springer Publishing Company, Incorporated.

[5] Van Der Aalst, W. M., Ter Hofstede, A. H. y Weske, M. (2003, Junio). "Business process management: A survey". En: *International conference on business process management* (pp. 1-12). Springer Berlin Heidelberg.

[6] Eid-Sabbagh, R. H., Dijkman, R., y Weske, M. (2012, Septiembre). "Business process architecture: use and correctness". En: *International Conference on Business Process Management* (pp. 65-81). Springer Berlin Heidelberg.

[7] Perez, M., Ferreyra, J. P., Verino, C., y Cocconi, D. (2017, Agosto). "Definición de una arquitectura de procesos utilizando la metodología BPTrends para la aplicación del ciclo de vida BPM". En: *XIX Workshop de Investigadores en Ciencias de la Computación* (WICC 2017, ITBA, Buenos Aires).

[8] Harmon, P. (2014). "*Business process change*", 3era. ed. Morgan Kaufmann.

Proceso de auditoría para micro, pequeñas y medianas organizaciones

Raúl Moralejo^{1,2}, Franco Catena², Matías Perez¹, Mara Rovero¹

¹ Escuela de Tecnología de la Información y las Comunicaciones

Departamento de Básicas y Tecnológicas

UNdeC - Universidad Nacional de Chilecito

romoralejo@gmail.com, mataguper20@gmail.com, mararovero@gmail.com.

² GridTICs – Grupo en Tecnologías de la Información y las Comunicaciones

Facultad Regional Mendoza

Universidad Tecnológica Nacional

romoralejo@gmail.com, francocatena@gmail.com

RESUMEN

Antes de profundizar en definiciones y en características de la auditoría, nos gustaría compartir una historia que habla sobre sus orígenes. Cierta o falsa, como pueden ser muchas de las historias del mundo, ésta en particular presenta mucha coherencia con la evolución que ha experimentado la auditoría y con los vicios que la han perseguido y la persiguen desde hace muchos años.

En los tiempos de la colonia británica en Norte América y al igual que todas las colonias europeas en el resto del mundo, los ingleses veían al nuevo continente como una fuente casi inagotable de riqueza que utilizaban para alimentar sus arcas y financiar nuevas aventuras y conquistas alrededor del globo.

Muchos ingleses de la comunidad de “inversionistas” en América, tenían fuertes intereses económicos en Estados Unidos, pero en aquellos tiempos no existían medios de transporte ni de comunicación como con los que contamos hoy en día. Esto presentaba el problema de que las operaciones económicas fueran manipuladas en perjuicio de los dueños británicos.

Considerando esto, se pensó que sería conveniente contar con alguna persona o personas de confianza que mantuviesen al tanto de los acontecimientos a los dueños “legítimos” de la riqueza que se generaba en América y que debería ser reportada y enviada al Reino Británico.

Estas personas fueron los primeros auditores, de hecho la palabra auditor tiene su origen en la palabra inglesa *Auditory*, que es el nombre de un nervio del aparato auditivo, llamado también “octavo nervio”. Este nervio, adicionalmente a su función auditiva, es el que permite mantener el sentido de equilibrio al ser humano.

Haciendo un paralelismo, podemos ver la relación que tiene este nervio con la función original del auditor y su primera encomienda, es decir informar sobre cualquier irregularidad en las operaciones y mantener un balance adecuado de las operaciones.

Este antecedente explica el por qué los auditores han sido “catalogados” a través del tiempo como personajes siniestros que se dedican a identificar todo lo que esté mal,

para denunciarlo y alertar a quien deba ser alertado.

Hoy en día, pensamos al auditor, como un elemento imprescindible para una sana operación de las empresas. Su rol ha pasado de ser un detector de problemas, a un identificador de oportunidades y emisor de propuestas de valor [1] [2].

Este proyecto pretende entender, analizar, y evaluar modelos de auditoría y generar buenas prácticas de control interno para ser aplicadas en el proceso de auditoría en micro, medianas y pequeñas organizaciones públicas y privadas usando como soporte el software mawidabp [3] desde una perspectiva CTS (Ciencia, Tecnología y Sociedad) [4].

Palabras clave:

Proceso de auditoría, control interno, modelos de auditoría, software mawidabp, software de código abierto, perspectiva CTS.

CONTEXTO

El presente trabajo está inserto en el proyecto de investigación presentado en la Universidad Nacional de Chilecito, Departamento de Básicas y Tecnológicas. Escuela de Tecnología de la Información y las Comunicaciones, la Cátedra Auditoría de Sistemas de la Carrera Ingeniería en Sistemas/Licenciado en Sistemas de la Universidad Nacional de Chilecito, y en la Universidad Tecnológica Nacional, Facultad Regional Mendoza, Grupo GridTICs, y la participación de la microempresa de software Cirope S.A. [5], quién brinda el software mawidabp que se utiliza como soporte para implementar las buenas prácticas. El software mawidabp se está utilizando en el área de auditoría interna de la Universidad Nacional

de Chilecito para llevar adelante el proceso de auditoría.

1. INTRODUCCION

Un auditor, hoy en día, no puede concebirse a sí mismo simplemente como el responsable de identificar riesgos en la operación de una empresa pública o privada, aunque ciertamente la identificación de riesgos sigue siendo una parte importante de sus actividades, su compromiso profesional va más allá de fungir como un mecanismo detectivo [1] [2].

Auditoría:

Podemos entender a la auditoría como la disciplina que mediante técnicas y procedimientos aplicados a una organización por personas independientes a la operación de la misma, evalúa el cumplimiento de los objetivos institucionales, emite una opinión al respecto y efectúa recomendaciones para mejorar el nivel de cumplimiento de dichos objetivos[1].

Proceso de auditoría:

La auditoría desarrolla las actividades por medio del siguiente proceso:

Planear la auditoría

Aún cuando la planeación se presenta como la primera actividad de la auditoría, en realidad el auditor debe haber recibido previamente a la planeación, un requerimiento para desarrollar un trabajo de auditoría, en este requerimiento se deben especificar los objetivos del trabajo y las condiciones generales del mismo. Con esta información, el auditor debe proceder a obtener un conocimiento general de la empresa que va a ser auditada, sus características de negocio, su infraestructura tecnológica, sus sistemas de información, sus áreas de riesgo, sus objetivos

estratégicos y cualquier asunto de interés específico sobre la auditoría a realizar.

Con esta información, el auditor debe realizar un trabajo de planeación en el que determine el tipo de procedimientos de revisión que deberá emplear, el personal responsable del desarrollo de las actividades y las fechas y duración aproximada del trabajo [1].

Analizar y evaluar el control interno

Con la información contenida en el plan de auditoría, el auditor preparará programas de trabajo específicos que le permitan obtener información sobre los procedimientos de control (elementos de administración) que la empresa tiene establecidos para apoyar sus objetivos de negocio. Esta información es obtenida mediante entrevistas, observación o inspección documental y debe ser documentada para servir como referencia y como fuente de consulta.

La información obtenida deberá ser empleada para fundamentar un análisis sobre la probable efectividad y eficiencia del sistema de control para lograr sus objetivos, cualesquiera que estos sean (integridad de la información, salvaguarda de activos, continuidad de operaciones, imagen, posicionamiento competitivo, etc.).

Aquí existe una premisa: Entre mejor sea el sistema de control, mayor será la probabilidad de que los objetivos sean alcanzados en forma satisfactoria [1].

Aplicar pruebas de auditoría

Con base en el análisis del control interno de la empresa, el auditor puede optar por aplicar dos tipos de pruebas. Las primeras se conocen como “pruebas de cumplimiento” y permiten verificar la efectividad de los procedimientos de control es decir, son técnicas de prueba que evalúan los “procesos de trabajo” existentes

en la empresa. Los resultados de estas pruebas podrán ratificar o rectificar el juicio preliminar sobre lo adecuado del control y sentar las bases para que el auditor aplique un segundo tipo de pruebas, estas últimas llamadas “pruebas sustantivas”. Las pruebas sustantivas están orientadas más bien a los “productos” de los procesos de trabajo y no a los procesos en sí. Los resultados de los pruebas sustantivas generalmente se expresan en información [1].

Informar sobre los resultados de la auditoría

Una vez realizadas las pruebas y analizados los resultados, el auditor debe preparar un informe sobre su trabajo, los resultados del mismo, las conclusiones correspondientes y las sugerencias que el auditor presente para mejorar las deficiencias encontradas. El auditor debe indicar claramente cuál es el nivel de involucramiento y responsabilidad en el trabajo realizado.

Dependiendo de los resultados y objetivos de su revisión, el auditor puede abstenerse de presentar una opinión cuando considere que no tiene elementos suficientes para sustentarla o puede presentar una opinión negativa, es decir, informar específicamente que los objetivos evaluados no han sido alcanzados en forma satisfactoria. El auditor también debe informar si se presentaron limitaciones importantes que evitaran que realizara su trabajo de acuerdo con las normas establecidas por su profesión [1].

Efectuar su seguimiento

Típicamente, una revisión arroja deficiencias o debilidades de control en la empresa, las que a su vez pueden representar oportunidades de mejora. Con base en estas oportunidades, el auditor emite

recomendaciones que se convierten en compromisos para los responsables de las áreas auditadas.

El auditor deberá efectuar revisiones de seguimiento para evaluar el nivel de cumplimiento de dichos compromisos o el impacto que esto pueda tener para la empresa. Los resultados de este tipo de trabajos también se deberán informar de manera adecuada a quién corresponda [1].

Supervisar el trabajo

En todo momento el auditor deberá ejercer una adecuada supervisión del trabajo, especialmente de aquel realizado por personal con menor experiencia, ya que en última instancia, el auditor será el principal responsable de su trabajo ante cualquier persona que lo utilice [1].

Control interno:

El control interno es el conjunto de elementos de administración que una empresa establece en forma coordinada para lograr sus objetivos institucionales y cumplir con su misión, es decir, cumplir con su razón de existir como empresa [1].

Modelos de auditoría:

COBIT (Control Objectives for Information and related Tecnologyes) es un modelo para auditar la gestión y control de los sistemas de información y tecnología, orientado a todos los sectores de una organización, es decir, administradores IT, usuario y por supuesto, los auditores involucrados en el proceso. COBIT es un modelo de evaluación y monitoreo que enfatiza en el control de negocios y la seguridad de IT y que abarca controles específicos de IT desde una perspectiva de negocios.

COSO (Committee of Sponsoring Organizations of Treadway Commission) se crea en Estados Unidos con la finalidad de identificar los factores que originan la emisión falsa o fraudulenta de reportes financieros.

El Informe COSO es un documento que contiene las principales directivas para la implantación, gestión y control de un sistema de Control Interno. Existen en la actualidad tres versiones del Informe, COSO I publicado en 1992, COSO II en el 2004 y la última versión publicada en el 2013 COSO III.

ITIL (Biblioteca de Infraestructura de TI), es un marco de trabajo de las mejores prácticas destinadas a facilitar la entrega de servicios de TI para ayudar a las organizaciones a lograr calidad y eficiencia en las operaciones de TI.

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

Los ejes principales son el proceso de auditoría, el control interno, modelos de auditoría como COBIT [6], COSO [7], ITIL [8], el uso del software mawidabp [3] para soporte del proceso de auditoría y la perspectiva CTS [4].

3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo principal del proyecto es generar buenas prácticas a partir de la investigación de los modelos de auditoría (COBIT, COSO, ITIL) para dar soporte al proceso de auditoría para aplicar en micro, pequeñas y medianas organizaciones públicas y privadas usando como soporte el software mawidabp [3] desarrollado con software de código abierto y una perspectiva CTS [4].

A la fecha, se han generado buenas prácticas de la investigación realizada sobre los

modelos COBIT, COSO e ITIL. Las mismas han sido cargadas en el software mawidabp para ser usadas por parte del área de auditoría interna de la Universidad Nacional de Chilecito. La misma está usando estas buenas prácticas por medio del software mawidabp a nivel de pruebas desde octubre de 2017 y en producción desde enero de 2018.

El objetivo hacia futuro es seguir investigando en detalle los modelos COBIT, COSO e ITIL para generar más buenas prácticas, las cuales serán cargadas en el software mawidabp como base del control interno, investigar en el proceso de auditoría desde de una perspectiva CTS, avanzar en la incorporación de otras áreas de auditoría interna de Universidades Nacionales de Argentina, y micro, pequeñas y medianas organizaciones de países iberoamericanos.

Tecnología de información utilizada para el desarrollo del software mawidabp

Sistemas operativos soportados

- Ubuntu 16.04, Debian 9.4, RedHat 7.4, y SLES 12 SP3

Software de base

- Interprete de Ruby MRI (versión mínima 2.5).
- Ruby on Rails (versión mínima 5.1).
- Servidor de BBDD PostgreSQL (versión mínima 10.1).
- Servidor web: Nginx (versión mínima 1.12).
- Servidor de aplicaciones Unicorn (versión mínima 5.4).
- Servidor NoSQL Redis (versión mínima 2.8).

4. FORMACION DE RECURSOS HUMANOS

Uno de los principales objetivos del proyecto es la capacitación de los recursos humanos, posibilitar la colaboración inter-institucional y la ejecución de tareas entre grupos de I+D y la microempresa de software Cirope S.A.

Para lograr estos objetivos se dispone del siguiente personal: 1 Investigador formado, 2 Investigadores de apoyo, y personal de la microempresa de software Cirope S.A.

Adicionalmente se realizarán:

Dictado de cursos, seminarios y conferencias.
Promoción, coordinación y asistencia técnica de tesinas para alumnos de la carrera de grado de ingeniería y licenciatura en sistemas de la UNdeC.

Presentación de trabajos en congresos y reuniones científicas/técnicas.

5. BIBLIOGRAFIA

- [1] Reingeniería de la Auditoría Informática. Gustavo Adolfo Solís Montes. 1999.
- [2] Auditoría en Informática. Un enfoque Metodológico. Enrique Hernández Hernández. 1997.
- [3] <https://www.mawidabp.com/>
- [4] Innovación en la Enseñanza de la Tecnología. Perspectiva CTS (Ciencia, Tecnología y Sociedad). Editorial Académica Española. Moralejo Raúl. 2012.
- [5] <https://www.cirope.com/>
- [6] ISACA. (Junio de 2016). Cursos-COBIT. Obtenido de Cursos-COBIT: <http://www.isaca.org/chapters7/Madrid/Event s/Formacion/Pages/Cursos-COBIT.aspx>
- [7] http://pkfperu.com/wp-content/uploads/2016/05/Interpretando_16.pdf
- [8] <https://www.itil.org.uk/>

Aproximaciones para el desarrollo multiplataforma de Aplicaciones Móviles

Marta C. Fennema, Susana I. Herrera, Rosa A. Palavecino, Pablo J. Najjar Ruiz, Paola D. Budán, Federico Rosenzvaig, Álvaro J. Carranza, Emmanuel Saavedra

Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero
{rosyppg, sherrera, pbudan, frosenzvaig@unse.edu.ar, najjarpablo@yahoo.com.ar}

Resumen

Esta línea de acción está incluida en el proyecto de investigación denominado “*Computación Móvil: desarrollo de aplicaciones y análisis forense*”. El mismo, está abocado al estudio de la generación de aplicaciones eficiente, evaluando los parámetros tiempo de desarrollo y cantidad de desarrolladores afectados. Para ello, se viene estudiado el uso de diversos *frameworks* generadores de aplicaciones móviles multiplataforma, a partir de un único proyecto web. Específicamente, esta línea de investigación enmarcada en la Ingeniería de Software, e indaga sobre el desarrollo eficiente de aplicaciones móviles multiplataforma que requieren acceso al hardware del dispositivo para funcionalidades con Realidad Aumentada basada en marcadores y objetos 3D. Los avances obtenidos se aplican en las áreas de salud, forense, y del *m-learning* principalmente.

Palabras clave: Desarrollo de aplicaciones móviles multiplataforma, Realidad Aumentada, marcadores, objetos 3D, *m-learning*.

1 Contexto

Esta investigación se lleva a cabo en el marco del proyecto denominado “*Computación Móvil: desarrollo de aplicaciones y análisis forense*”, del Instituto de Investigaciones en Informática y Sistemas de Información (IISI) de la Universidad Nacional de Santiago del Estero (UNSE). El mismo se está ejecutando desde el año 2017 y su finalización está prevista para fines de 2018. Es una continuación del proyecto “Optimización de la calidad de los Sistemas Móviles mediante la implementación de nuevas arquitecturas, Realidad Aumentada (RA), técnicas de visualización y redes móviles Ad-Hoc. Aplicaciones en *m-learning* y en gestión del conocimiento”, implementado entre los años 2012 y 2016.

El equipo de investigación del proyecto está conformado por docentes del Instituto de Investigaciones en Informática de la Facultad de Ciencias Exactas de la Universidad Nacional de Santiago del Estero (UNSE). Además, el proyecto cuenta con el asesoramiento de investigadores de Institutos de Investigación de la Facultad de Informática de la Universidad Nacional de La Plata.

El contexto de desarrollo de esta investigación está dado por las condiciones del contexto móvil a nivel mundial, donde constantemente están surgiendo nuevas herramientas tanto para generar código general, como bibliotecas específicas, como las que sustentan la incorporación de la RA.

2 Introducción

El equipo de investigadores atiende la problemática relacionada con el desarrollo eficiente de aplicaciones móviles multiplataforma (Fennema et al., 2017). Es decir, aplicaciones móviles que puedan ser ejecutadas en los Sistemas Operativos (SO) más usados en Argentina: Android y iOS (StatCounter Global Stats, 2015)). Se trabaja con aplicaciones complejas que requieren acceder a los recursos del hardware del dispositivo móvil (cámara, GPS, acelerómetro). Se busca un desarrollo eficiente en función de la optimización del tiempo de desarrollo y de la cantidad de desarrolladores afectados. Para ello, se han estudiado el uso de diversos *frameworks* que permiten generar aplicaciones móviles híbridas y nativas para diversos SO a partir de un único proyecto web.

En relación a su arquitectura, un sistema móvil generalmente se basa en una aplicación cliente que se conecta a un servidor de aplicaciones que se encuentra en Internet (Challiol, 2012). Este servidor, a su vez, utiliza los servicios de un proveedor de ubicación y la información provista por diversos puntos de interés. Más profundamente, las arquitecturas

alternativas que se pueden adoptar en el momento de desarrollar una aplicación móvil fueron estudiadas y analizadas en (Najar et al., 2014). Ellas son:

- **Arquitectura Web:** la aplicación reside completamente en el servidor web y es accedido a través del browser del teléfono. Puede ser accedida por la mayoría de los dispositivos. Se dificulta la interacción con los periféricos del dispositivo, por ello se debe utilizar otras tecnologías como por ejemplo el uso de códigos QR para la determinación del contexto de localización.
- **Arquitectura Cliente-Servidor o híbrida:** parte de la aplicación en el cliente y parte en el servidor. La aplicación cliente consulta la aplicación servidor a través del uso de Web Services. Se tiene en cuenta características del SO y características del dispositivo donde se instalará. Permite interactuar con los periféricos del dispositivo (cámaras, bluetooth, GPS) que permitirán determinar los datos de contexto en el que se encuentra el móvil. Se llaman aplicaciones nativas, desarrolladas para cada SO móvil.
- **Arquitectura Cliente:** la aplicación reside completamente en el cliente, la información o las bases de datos residen en el dispositivo. También son consideradas aplicaciones nativas.

En la presente investigación, interesan aplicaciones que permiten la interacción entre los usuarios. Por lo tanto, no abarca las arquitecturas clientes.

Actualmente el desarrollo de aplicaciones móviles tiende al desarrollo multiplataforma. Según Delia (2017), las aplicaciones multiplataforma pueden clasificarse en: aplicaciones web móviles, híbridas, interpretadas y generadas por compilación cruzada.

- **Aplicaciones Web móviles:** diseñadas para ejecutarse dentro de un navegador, se desarrollan con tecnología web estándar (HTML, CSS y JavaScript), no necesitan adecuarse a ningún entorno operativo, son independientes de la plataforma. Sin embargo, sus tiempos de respuesta decaen debido a la interacción cliente-servidor, son menos atractivas que las aplicaciones nativas ya que no se encuentran instaladas en el dispositivo, lo que implica acceder previamente a un navegador. Además, las restricciones de seguridad impuestas a la ejecución de código por medio de un navegador, limitan el acceso a todas las capacidades del dispositivo.
- **Aplicaciones híbridas:** utilizan tecnologías web (HTML, Javascript y CSS) pero no son

ejecutadas por un navegador. En su lugar, se ejecutan en un contenedor web (webview), como parte de una aplicación nativa, la cual está instalada en el dispositivo móvil. Las aplicaciones híbridas permiten la reutilización de código en las distintas plataformas, el acceso al hardware del dispositivo, y la distribución a través de las tiendas de aplicaciones. Respecto de las nativas, poseen las siguientes desventajas: i) la experiencia de usuario se ve perjudicada al no utilizar componentes nativos en la interfaz, y ii) la ejecución se ve ralentizada por la carga asociada al contenedor web. Existe una diversidad de *frameworks* que permiten desarrollar aplicaciones híbridas: *PhoneGap*¹, *Apache Cordova* (Wargo, 2015), *CocoonJS*, *Ionic*², *SenchaTouch*.

- **Aplicaciones interpretadas:** son implementadas utilizando un lenguaje base, el cual se traduce en su mayor parte a código nativo, mientras el resto es interpretado en tiempo de ejecución. Estas aplicaciones son implementadas de manera independiente de las plataformas utilizando diversas tecnologías y lenguajes, tales como *Javascript*, *Java*, *Ruby* y *XML*, entre otros. Una de las principales ventajas de este tipo de aplicaciones es que se obtienen interfaces de usuario totalmente nativas. Sin embargo, los desarrolladores experimentan una dependencia total con el entorno de desarrollo elegido. Los *frameworks* más conocidos para generar estas aplicaciones son *Appcelerator Titanium* y *NativeScript*.
- **Aplicaciones generadas por compilación cruzada** Estas aplicaciones se compilan de manera nativa creando una versión específica de alto rendimiento para cada plataforma destino. Ejemplos de entornos de desarrollo para generar aplicaciones por compilación cruzada son *Applause*, *Xamarin* (Hermes, 2015), *QT*, *Embarcadero Delphi 10 Seattle* y *RubyMotion*.

Luego de haber realizado un conjunto de prácticas de prueba, en esta línea de investigación se decide comparar *QT*³ y *Xamarin* (Hermes, 2015). El primero, permite generar aplicaciones nativas para varias plataformas y tiene un amplio soporte al desarrollador. El API de la biblioteca cuenta con métodos para acceder a bases de datos mediante SQL, así como uso de XML, gestión de hilos, soporte de red, una API multiplataforma unificada para la manipulación de archivos y una

¹ <https://www.genbetadev.com/frameworks/phonegap>

² <https://ionicframework.com/>

³ <https://www1.qt.io/developers/>

gran variedad de métodos para el manejo de ficheros, además de estructuras de datos tradicionales. Por su parte, *Xamarin* es una plataforma para el desarrollo de aplicaciones móviles nativas para Android, iOS y Windows utilizando el lenguaje C#. Las interfaces de usuario de las aplicaciones escritas en C#/*Xamarin* son nativas, lo que produce una mejora en el rendimiento en tiempo de ejecución. Las aplicaciones escritas en C#/*Xamarin* permiten un enfoque de desarrollo multiplataforma donde se comparte la codificación de lógica de negocio. Sin embargo, las interfaces de usuario deben ser programadas independientemente para cada plataforma de destino.

Tal como sostiene Delía (2017), la elección de qué *framework* emplear depende del tipo de aplicación que se quiere lograr y del SO del dispositivo en el cual se ejecutará la misma.

En cuanto a la RA, es una tecnología que complementa la percepción e interacción con el mundo real, ya que proporciona un entorno real aumentado con información adicional generada por la computadora o algún dispositivo móvil. Posibilita el desarrollo de aplicaciones interactivas que combinan la realidad con información sintética, tal como imágenes 3D, sonidos, videos, textos, sensaciones táctiles, en tiempo real, y de acuerdo al punto de vista de quien está observando la escena (Salazar Mesía et al., 2015). Se distinguen dos tipos básicos de RA:

- Geolocalizada, la cual a través de sensores indican el posicionamiento del dispositivo móvil. Con solo sostener el móvil en nuestro campo de visión, podemos visualizar puntos de interés cercanos gracias a la información tomada a través de varios sensores móvil, entre ellos el GPS y brújula de orientación.
- Basada en marcadores, que reconocen patrones de activadores de información, como puede ser un código de barras, QR o un símbolo, en la imagen de video que se recibe desde una cámara. Cuando se reconoce un patrón en particular, en su posición, se superpone una imagen digital en la pantalla. Este es el primer tipo de RA, que tuvo sus orígenes en algo muy sencillo: etiquetas.

Los avances en el campo de la Ingeniería del Software se aplican sobre dos dominios relevantes para la sociedad: salud (Córdoba, 2016) y educación (Herrera et al., 2017; Herrera & Sanz, 2014; Herrera et al., 2014). Además, actualmente se está trabajando para lograr una aplicación para *m-learning* que le permita al alumno desarrollar la capacidad de representar

situaciones de la vida cotidiana mediante sistemas de ecuaciones lineales. La misma se denomina AlgeRA y tiene por objetivos específicos: a) ejemplificar con RA sistemas de ecuaciones lineales que representan situaciones de la vida cotidiana, b) modelizar el comportamiento de una Red de tránsito vehicular usando Sistemas de Ecuaciones Lineales. La primera funcionalidad constituye una “Ejemplificación” mientras que, la segunda, es un “Trabajo de campo para la modelización”.

3 Líneas de investigación y desarrollo

La principal línea de investigación de esta propuesta consiste en el estudio de métodos, técnicas y herramientas que permitan el desarrollo de sistemas móviles multiplataforma; lo cual significa indagar sobre:

- Herramientas para la generación automática de aplicaciones móviles nativas que interactúan con hardware del dispositivo,
- Métodos ágiles para el desarrollo de aplicaciones móviles multiplataforma.

En esta propuesta se pretende investigar sobre plataformas o entornos de desarrollo de aplicaciones móviles que permite generar aplicaciones nativas que puedan ser optimizadas mediante la modificación del código.

En cuanto al dominio de aplicación, se continuará trabajando en Educación y Salud, desarrollando aplicaciones para *m-learning*, para rehabilitación de personas con discapacidad cerebral, y para el dominio de la informática forense.

En lo que respecta a *m-learning*, el trabajo se está enfocando al desarrollo de AlgeRA, lo cual requiere investigar sobre la interacción de *frameworks* específicos de RA con *Xamarin*, y el uso de herramientas de *sincronicidad* que permitan obtener los recursos para el *trabajo de campo*.

4 Objetivos y resultados

El objetivo general que guía esta investigación aplicada es:

Contribuir al progreso del campo de la Computación Móvil mediante: a) la optimización del desarrollo y mantenimiento de aplicaciones móviles nativas multiplataforma, b) el estudio de herramientas de RA, y c) el análisis forense de dispositivos móviles.

Los objetivos específicos son:

- a) Analizar herramientas (libres y propietarias) de generación automática de aplicaciones nativas para múltiples SO móviles.
- b) Definir métricas para evaluar la capacidad de mantenimiento de aplicaciones móviles con herramientas de generación automática.
- c) Comparar herramientas para el desarrollo de aplicaciones móviles con RA.
- d) Definir un método ágil apropiado para el desarrollo de aplicaciones móviles nativas y multiplataforma, basadas en generación automática de código.
- e) Determinar la calidad de aplicaciones móviles construidas con las diversas herramientas analizadas, usando prototipos construidos con métodos ágiles y las métricas definidas.

Esta línea de investigación ha transitado su primera etapa de un año, abocada principalmente a la comparación de las herramientas de desarrollo multiplataforma Xamarin y QT. Además, dados los requerimientos iniciales de AlgeRA, se han comparado las herramientas nombradas en cuanto a la posibilidad de incorporar herramientas para la RA.

Las publicaciones obtenidas hasta la fecha son:

- Fennema, M., Herrera, S., Palavecino, R., Budán, P., Rosenzvaig, F., Najar Ruiz, P., Carranza, A., Saavedra, E. (2017). Aproximaciones para el desarrollo multiplataforma y mantenimiento de Aplicaciones Móviles. Workshop en Investigación en Ciencias de la Computación, WICC. Instituto Tecnológico Buenos Aires. Buenos Aires. ISBN 978-987-42-5143-5.
- Susana I. Herrera, Paola D. Budán, Federico Rosenzvaig, Marilena Maldonado, Gabriela Suárez, Emmanuel Saavedra, Pablo J. Najar Ruiz, María I. Morales. Desarrollo de Aplicaciones Multiplataforma para M-learning con Realidad Aumentada. ADNTIIC 2017. Huerta Grande. Córdoba. Argentina.

Este año se prevé cumplimentar los objetivos b), d) y e) propuestos en esta línea de acción. En relación al objetivo e) se dispone de base la evaluación realizada por Herrera et al. (2013).

5 Formación de recursos humanos

La Directora del proyecto pertenece al Instituto de Investigación en Informática y Sistemas de Información de la UNSE. La Codirectora

pertenece a la Universidad del Norte Santo Tomás de Aquino.

Actualmente, se están realizando trabajos finales de LSI sobre: realidad aumentada, desarrollo híbrido de aplicaciones, métodos de desarrollo de sistemas móviles, sistemas sensibles al contexto temporal, herramientas para medición de magnitudes físicas.

Se tiene previsto incorporar más alumnos de grado en los próximos años, a efectos de brindarles un marco apropiado para la elaboración de sus trabajos finales.

Referencias

1. Challiol, C. (2012). Apuntes de Curso de Posgrado sobre Computación Móvil. Curso dictado en Universidad Nacional de Santiago del Estero. Secretaría de Extensión de la Facultad de Ciencias Exactas y Tecnologías.
2. Córdoba, Melisa N. V.; Najar, Pablo; Budán Paola Daniela. (2016). Sistema Alternativo de Comunicación para Niños con Parálisis Cerebral Infantil. Argentina. San Salvador de Jujuy. 2016. Libro. Artículo Completo. Jornada. XI Jornadas de Ciencia y Tecnología de Facultades de Ingeniería del NOA. CODINOA.
3. Delía, L. (2017). Desarrollo de aplicaciones móviles multiplataforma. Trabajo Final de Especialista en Ingeniería Web. Facultad de Informática, Universidad Nacional de La Plata.
4. Fennema, M., Herrera, S., Palavecino, R., Budán, P., Rosenzvaig, F., Najar Ruiz, P., Carranza, A., Saavedra, E. (2017). Aproximaciones para el desarrollo multiplataforma y mantenimiento de Aplicaciones Móviles. Workshop en Investigación en Ciencias de la Computación, WICC. Instituto Tecnológico Buenos Aires. Buenos Aires. ISBN 978-987-42-5143-5
5. Hermes, D. (2015). Xamarin Mobile Application Development: Cross-Platform C# and Xamarin.Forms Fundamentals. Ed. Apress. ISBN 978-1484202159.
6. Herrera, S., Najar, P., Palavecino, R., Goñi, J. (2013) Evaluación de la calidad en aplicaciones móviles. Investigaciones en Facultades de Ingeniería del NOA. Santiago del Estero: Universidad Nacional de Santiago del Estero. vol. n°. p103 - 104. issn 1853-7871.

7. Herrera, S. I. & Sanz, C. (2014). Collaborative m-learning practice using Educ-Mobile. International Conference on Collaboration Technologies and Systems (CTS). Ed. IEEE. Pp 363 – 370. ISBN: 978-1-4799-5157-4. Minneapolis, MN, USA.
8. Herrera, S., Morales, M.I., Fennema, M. C. y Sanz, C.V. (2014). Aprendizaje basado en dispositivos móviles. Experiencias en la Universidad Nacional de Santiago del Estero. ISBN 978-987-1676-18-7. Ed. EDUNSE. Santiago del Estero.
9. Herrera, S., Palavecino, R., Sanz, C., Carranza, J. (2017). Aprendizaje de Estructuras de Datos mediante m-learning. I Simpósio Ibero-Americano de Tecnologias Educacionais. SITED 2017. ISSN 2594-388X. Universidade Federal de Santa Catarina. Araranguá, SC, Brasil.
10. Ionic. (2016). Disponible en: <http://ionicframework.com/docs/guide/preface.html>. Fecha de acceso: 13/08/2016.
11. Lliteras, A., Challiol, C., Gordillo S. (2012) Juegos Educativos Móviles Basados en Posicionamiento: Una Guía para su Conceptualización. 13th Argentine Symposium on Software Engineering, 41 JAIIO, La Plata.
12. Lliteras, A., Challiol, C., Mostaccio, C., Gordillo S. (2011). Representaciones enriquecidas para la navegación indoor-outdoor en aplicaciones móviles. XVII Congreso Argentino de Ciencias de la Computación. La Plata
13. Najar, P., Ledesma, E., Rocabado, S., Herrera, S. y Palavecino, R. (2014). Eficiencia de aplicaciones móviles según su arquitectura. CACIC 2014. ISBN 978-987-3806-05-6. La Matanza, Buenos Aires.
14. PhoneGap. (2013). A guid to building cross-platform apps using the W3C standards based Cordova/PhoneGap framework. Packt Publishing.
15. Salazar Mesía, N, Gorga, G., Sanz, C. (2015). Plan de evaluación del material educativo digital EPRA. Propuesta de indagación sobre la motivación intrínseca. XXI Congreso Argentino de Ciencias de la Computación, Buenos Aires (UNNOBA) Sede Junín.
16. StatCounter Global Stats (2015). Ranking de los 8 Sistemas Operativos Móviles más usados en Argentina. Disponible en <http://gs.statcounter.com/#mobile_os-AR-monthly-201401-201501-bar>. Fecha de consulta: 18/04/2015.
17. Wargo, J. (2015). Apache Cordova 4 Programming. Addison-Wesley Professional.

Aplicación móvil multiplataforma sensible al contexto para el turismo religioso en Santiago del Estero

Susana I. Herrera¹, María M. Clusella¹, Silvia Sánchez Zuaín¹, Sergio H. Rocabado Moreno², David Cheein¹, Fernando Leturia¹, Stefano Trejo¹

¹*Facultad de Ciencias para la Innovación y el Desarrollo, Universidad Católica de Santiago del Estero*

²*Facultad de Ciencias Exactas, Universidad Nacional de Salta*
suiherrera@gmail.com, mercedesclusella@gmail.com, silvisanzu@gmail.com,
rocabado@di.unsa.edu.ar, josedavidcheein@gmail.com, fleturia@gmail.com, stefanots@gmail.com

Resumen

Es una investigación aplicada que involucra estudios en el campo de la Ingeniería del Software para desarrollar una aplicación móvil multiplataforma sensible al contexto para el Turismo Religioso en la Provincia de Santiago del Estero, Argentina. La aplicación se basa en un modelo sistémico que permitirá a los diversos actores del turismo la gestión dinámica del contenido. La aplicación brindará al turista información en línea del entorno, sea cual fuere el SO de su dispositivo móvil; al resto de los actores del turismo (proveedores, guías, administradores de eventos e instalaciones turísticas, etc.) la aplicación les permitirá actualizar en línea la información de contexto. La investigación involucra el análisis comparativo de herramientas de desarrollo híbrido de aplicaciones móviles (PhoneGap, Ionic, Apache Córdoba), según criterios de reuso, versionado y manejo de bibliotecas de GPS para la sensibilidad al contexto. Además, se estudia la conectividad inalámbrica móvil en zonas rurales, abordando mecanismos alternativos de conectividad que aseguren el uso efectivo de las aplicaciones móviles en contextos rurales aislados (escasa conectividad y recursos energéticos limitados), típicos de los lugares donde se desarrollan importantes festividades religiosas santiagueñas.

Palabras clave: aplicaciones móviles multiplataforma, sensibilidad al contexto, turismo religioso, religiosidad en Santiago del Estero.

1 Contexto

Esta investigación se desarrolla con la aprobación y el financiamiento por la Secretaría de Ciencia y Tecnología de la Universidad Católica de Santiago del Estero (UCSE). Se inició en el mes de Julio de 2017 y su finalización está prevista para Junio de 2019.

El equipo de investigación del proyecto está conformado por docentes 3 docentes investigadores de la Facultad de Ciencias para la Innovación y el Desarrollo (FCID) de la UCSE, 1 docente investigador de la Facultad de Ciencias Exactas de la Universidad Nacional de Santiago del Estero (UNSa), 2 alumnos de posgrado y 1 de grado de carreras de Informática de la FCID.

La investigación surgió a partir de un estudio de necesidades del medio organizado por UCSE y, a su vez, dando continuidad a investigaciones previas sobre turismo religioso desarrolladas por la misma universidad en años anteriores. Es decir, permitirá optimizar la gestión del turismo religioso en la provincia y, por lo tanto, contribuye al objetivo de la UCSE referido a aportar soluciones de calidad a problemáticas concretas de la sociedad santiagueña a través de la I+D+i.

La investigación se desarrolla en colaboración con el equipo de investigación en Computación Móvil de la Universidad Nacional de Santiago del Estero y con la Subsecretaría de Turismo de la Provincia de Santiago del Estero, con quienes está en proceso la firma de un convenio formal de colaboración. Asimismo los trabajos finales de posgrado cuentan con el asesoramiento de un investigador del Instituto de Investigación en Informática LIDI de la Universidad Nacional de La Plata.

2 Introducción

El turismo religioso en Santiago del Estero es una temática de análisis en respuesta a la gran cantidad de creyentes, eventos, festividades religiosas y a la vez a la búsqueda de nuevas tendencias y captación de demandas turísticas en la Provincia. Se caracteriza por convocar grandes multitudes a través de sus diferentes manifestaciones. Es la modalidad de viaje en la cual la fe, las creencias, la espiritualidad, la

devoción constituyen las motivaciones principales de este tipo de viajeros, sumado al deseo de descanso y curiosidad por conocer paisajes, manifestaciones, costumbres, lugares y personas.

Las nuevas tecnologías han modificado las pautas en la experiencia del turista en el lugar de destino, haciendo surgir un nuevo tipo de consumidor de turismo, el turista 3.0 o turista colaborativo. Este interactúa con su destino a través de dispositivos móviles y es asiduo a redes sociales e internet. Actualmente este tipo de dispositivos se han convertido en una herramienta clave, dado que hay más móviles que computadoras y las personas pasan muchas más horas con ellos que con cualquier otro dispositivo computacional. De hecho, los móviles rara vez se apagan y pocas veces se desconectan de la red. Los sistemas móviles constituyen uno de los pilares fundamentales en el fomento del turismo en muchos lugares del mundo. Son importantes los avances logrados por algunos integrantes del equipo de esta propuesta respecto a la mejora de la gestión del turismo usando móviles y otras TIC (Herrera, Najar Ruiz, Contreras, Fenema & Lara, 2013; Herrera, Gallo & Najar Ruiz, 2013; Herrera, Najar Ruiz, Ledesma & Rocabado, 2012); así como también los avances sobre la gestión de aspectos de la cultura usando TIC (Gallo, Palavecino & Herrera, 2014; Herrera, Zuain, Gallo & Avila, 2012; Herrera, Clusella, Mitre, Santillán & García, 2011; Herrera, Clusella, Luna, Mitre & Santillán, 2011).

La Computación Móvil es un área de la Informática que ha evolucionado vertiginosamente en los últimos años. Se basa en el estudio de los dispositivos móviles, redes móviles y aplicaciones móviles (Talukder et al., 2010). De acuerdo a Siegler (2008), se entiende por aplicación móvil o sistema móvil a un programa informático diseñado para ser ejecutado en teléfonos inteligentes (smartphones), tabletas y otros dispositivos móviles. Pueden responder a diversas arquitecturas: cliente, web e híbrida; son importantes los estudios sobre la calidad de sistemas móviles para turismo según su arquitectura realizados en esta región por Najar Ruiz, Ledesma, Rocabado, Herrera & Palavecino (2014). Poseen diversos sistemas operativos (SO) móviles entre los que se estacan Android de Google, iOS de Apple y Windows Phone de Microsoft. Dado que existe una gran variedad de dispositivos móviles en el mercado, es deseable que una aplicación sea multiplataforma, es decir, que sea compatible con cualquier SO, evitando así condicionar al usuario a utilizar una plataforma u otra.

Sin embargo, el desarrollo de aplicaciones multiplataforma (arquitectura web) implica un gran esfuerzo desde la Ingeniería del Software, ya que involucra la necesidad de desarrollar aplicaciones para cada SO, utilizando distintos lenguajes y distintos entornos (IDE). Esto, a su vez, implica: el uso de más recursos (desarrolladores y tiempo), menor posibilidad de reuso del software, mayor esfuerzo de actualización de software, aumento de posibilidades de error en el código y mayor esfuerzo en la gestión de configuración (Emmanouilidis et al., 2013). En los últimos tiempos, han surgido herramientas automáticas de desarrollo de aplicaciones móviles, entre ellas: PhoneGap (PhoneGap3, 2013), Ionic (Ionic, 2016), ApacheCórdoba (Wargo, 2015). El desarrollo de software usando este tipo de herramientas se conoce como desarrollo híbrido de aplicaciones; pero se desconoce la eficiencia de las mismas. Es por ello que, en esta investigación, se propone analizar y comparar las distintas herramientas de desarrollo híbrido de aplicaciones móviles, considerando los aspectos mencionados anteriormente en este párrafo. En relación al desarrollo de arquitectura web es importante mencionar, que algunos integrantes del equipo, ya vienen investigando en esta temática (Sanchez Zuain & Durán, 2016).

En particular, se incorporó a la investigación la clasificación de Delía (2017), que sostiene que las aplicaciones multiplataforma pueden clasificarse en: aplicaciones web móviles, híbridas, interpretadas y generadas por compilación cruzada.

En particular, los sistemas móviles sensibles al contexto proveen información en línea al usuario según la ubicación en la que se encuentra, brindándole diversos servicios, teniendo en cuenta, en algunos casos, su contexto social y sus preferencias. El desarrollo de este tipo de sistemas implica la necesidad de utilizar bibliotecas de funciones específicas para la interacción con el GPS del dispositivo. Actualmente, las aplicaciones de turismo utilizan estas funciones, según cada SO. Pero no existen estudios sobre la eficiencia del manejo de GPS en el desarrollo híbrido de aplicaciones. Esta problemática también será abordada en esta propuesta.

Por otra parte, para asegurar una buena experiencia de usuario en el uso de una aplicación móvil, se debe asegurar una conectividad inalámbrica de calidad. La provincia de Santiago del Estero se caracteriza por poseer zonas aisladas, con conectividad y energía eléctrica

escasa o nula. Por lo tanto, se deberá estudiar formas alternativas de conectividad móvil. En este sentido, algunos de los investigadores del equipo propuesto poseen también estudios previos en la temática (Rocabado, Cadena & Díaz, 2016; Rocabado, Sanchez, Herrera & Cadena, 2016; Rocabado, Cadena & Díaz, 2015; Rocabado, Herrera, Morales & Estellés, 2013; Herrera, Rocabado, Coronel & Campos, 2013). Esta investigación abordará mecanismos alternativos de conectividad para asegurar el uso efectivo de las aplicaciones móviles en contextos rurales aislados, típicos de los lugares donde se desarrollan importantes festividades religiosas santiagueñas.

Sintetizando, en esta propuesta se pretende realizar una investigación aplicada, que involucra el desarrollo de una aplicación móvil multiplataforma sensible al contexto que gestione contenido dinámico siguiendo el modelo propuesto por Herrera et al. (2013-a). La aplicación brindará al turista, sea cual fuere el SO de su dispositivo, información del entorno en línea. Mientras que al resto de los actores del turismo (proveedores, guías, administradores de eventos e instalaciones turísticas, etc.) permitirá actualizar en línea la información de contexto. La investigación contribuirá principalmente al campo de la Ingeniería del Software, ya que analizará la eficiencia de herramientas de desarrollo híbrido de aplicaciones móviles, según criterios de reuso, versionado y manejo de bibliotecas de GPS para la sensibilidad al contexto. Además, aportará avances en el estudio de la conectividad inalámbrica móvil en zonas rurales.

3 Líneas de investigación y desarrollo

Se trata de una investigación aplicada, desarrollada desde un enfoque cuantitativo, en el campo de la Informática.

la hipótesis que guiará la investigación propuesta es la siguiente:

El uso de herramientas de desarrollo híbrido permite desarrollar eficientemente aplicaciones móviles multiplataforma sensibles al contexto.

En la misma se identifican claramente dos variables que están relacionadas entre sí: por un lado, herramientas para el desarrollo híbrido y, por otro, el desarrollo eficiente de aplicaciones móviles multiplataforma sensibles al contexto.

La investigación constituye un estudio de caso, por lo tanto, no es necesario definir universo ni muestra. El caso constituye el turismo religioso en Santiago del Estero.

La investigación requiere la obtención de conocimientos del dominio de aplicación, el Turismo. Para ello, se trabajará con informantes calificados (especialistas de Turismo y responsables de la gestión del turismo en la provincia), aplicando las siguientes técnicas: entrevistas basadas en cuestionarios de preguntas abiertas, sesiones de trabajo grupal aplicando técnicas de creatividad. También se aplicará búsqueda web, análisis de documentos.

Para el desarrollo de prototipo se utilizará Mobile-D y técnicas de desarrollo basadas en UML.

4 Objetivos y resultados

Los objetivos generales son:

- Contribuir a la optimización del desarrollo híbrido de aplicaciones móviles, mediante el estudio de la eficiencia de diversas herramientas de generación de aplicaciones móviles multiplataforma.
- Desarrollar un prototipo de aplicación móvil multiplataforma sensible al contexto para el turismo religioso en la Provincia de Santiago del Estero.

Los objetivos específicos son:

- Definir indicadores que permitan analizar y comparar la eficiencia de las herramientas de desarrollo híbrido de aplicaciones móviles.
- Analizar el comportamiento de las diversas herramientas de desarrollo híbrido respecto a la gestión del GPS de dispositivos móviles con diferentes SO.
- Analizar las diferentes alternativas de conectividad móvil para aplicaciones de turismo en la provincia y determinar la/s tecnologías y configuraciones más convenientes.
- Adaptar el modelo de Herrera et al. (2013-a), para la aplicación móvil de turismo religioso de Santiago del Estero
- Diseñar el prototipo de aplicación móvil sensible al contexto para el turismo religioso de Santiago del Estero, utilizando un método ágil
- Desarrollar prototipos utilizando diversas herramientas para el desarrollo híbrido de aplicaciones móviles

- Comparar la eficiencia de las herramientas de desarrollo híbrido, según los prototipos implementados y los indicadores definidos previamente.

Esta investigación ha transitado sólo 8 meses desde su inicio, habiéndose obtenido los siguientes resultados parciales:

- Se realizó el análisis comparativo de 3 herramientas de desarrollo híbrido de aplicaciones móviles: ApacheCordova, Ionic y PhoneGap. Se definieron aspectos a considerar en la comparación. En función del análisis, se decidió desarrollar el prototipo usando ApacheCordova.
- Se realizó un prototipo usando ApacheCordova para mostrar el comportamiento de aplicaciones que usan GPS generadas desde este framework, usando simuladores de GPS y emuladores de móviles. Con ello se pudo corroborar la viabilidad del desarrollo de aplicaciones sensibles al contexto usando ApacheCordova.
- Se analizaron equipamientos alternativos para amplificación de recepción de Red de Datos. Se definieron las características de conectividad de los equipos: amplificadores compatibles no sólo con 4G sino también con 3G y 2G (dado que en las zonas rurales de Santiago del Estero no cuentan con antenas 4G). Se realizó la compra de repetidor/amplificador de señal de celular, diseñado para Banda celular 2G 850MHz - 3G 850MHz y 4G 1700MHz 2100MHz AWS, BAND 4. Se realizó la prueba del equipamiento usando el software Network Signal Info, en el asentamiento Sauce Bajada, Departamento Banda.
- Se realizó la revisión del modelo de turismo de Herrera et al. (2013-a), el cual prevé que los servidores con la información deben estar centralizado y a cargo de un ente gubernamental responsable del Turismo en la provincia. Se inició el análisis de la evaluación de factibilidad. Para ello, se realizó entrevista al Subsecretario de Turismo de la Provincia, periodista Ricardo Sosa, en diciembre 2017. Según el mismo, sería viable su implementación en dicho organismo.
- Usando Mobile-D, se inició el proceso de análisis de requisitos de la aplicación. El principal objetivo de la App sería: Definir una ruta de Turismo Religioso para la provincia de Santiago del Estero. Para ello se realizaron las siguientes actividades:

- Exploración de aplicaciones móviles de turismo religioso para revisar su comportamiento más apps de turismo de Santiago del Estero: Turinea, Santiago de Compostela, El llamador, Cañari, BA Iglesias, SDE TOUR, Santiago Turismo.
- Revisión de bibliografía y de fuentes de información de Turismo Religioso. Técnicas: entrevistas y exploración web. Las principales fuentes consultadas fueron: Proyecto de Investigación UCSE sobre Turismo Religioso en Santiago del Estero (Dra. Rossi), Revista Trazos (Lic. Mirta Caumo). Entrevistas a: coordinadora de Turismo de la Facultad de Cs Económicas UCSE y a ex directora de Turismo de la Municipalidad de SDE (Prof. Vaultet).
- Definición de informantes calificados para el relevamiento.
- Definición del instrumento para obtención de información calificada, entrevista basada en cuestionario abierto.

5 Formación de recursos humanos

La investigación contribuye a la formación en investigación de 2 alumnos de posgrado de la carrera Especialización en Sistemas Web de la UCSE. También contribuye a la formación en investigación de un alumno de grado, mediante su participación como ayudante estudiantil en la cátedra Metodología de la Investigación de la carrera Ingeniería en Informática.

Referencias

1. Delía, L. (2017). Desarrollo de aplicaciones móviles multiplataforma. Trabajo Final de Especialista en Ingeniería Web. Facultad de Informática, Universidad Nacional de La Plata.
2. Emmanouilidis, C., Koutsiamanis, R.A., Tasidou, A. (2013). Mobile guides: Taxonomy of architectures, context awareness, technologies and applications. *Journal of Network and Computer Applications*. Volume 36, Issue 1, January 2013, Pages 103–125. Elsevier.
3. GALLO, F.; PALAVECINO, R; HERRERA, S. 2014. Evaluación de la Usabilidad en sistemas E-Cultura. Actas del 2º Congreso Nacional de Ingeniería Informática/Sistemas de Información. San Luis: Universidad Nacional de San Luis, Facultad de Ciencias Físico-Matemáticas y Naturales, Departamento de Informática. vol. n°. p211 - 224. issn 2346-9927.

4. HERRERA, S.; NAJAR RUIZ, P.; CONTRERAS, N.; FENNEMA, C.; LARA, C. (2013-a). Marco Sistémico para el Desarrollo de Aplicaciones de m-Turismo. Brasil. Palmas, Tocantins. Libro. Artículo Completo. Congreso. Congresso Brasileiro de Sistemas: pensar a compreensão de sistemas. Universidade Federal do Tocantins.
5. HERRERA, S.; GALLO, F.; NAJAR RUIZ, P. (2013-b). Accesibilidad en aplicaciones móviles para turismo. Argentina. Córdoba. Libro. Artículo Completo. Congreso. Segundo Congreso Argentino de la Interacción-Persona Computador@, Telecomunicaciones, Informática e Información Científica IPCTIIC 2013. ALAIPO.
6. HERRERA, S.; ROCABADO, S.; CORONEL, J.; CAMPOS, M. (2013-c). Redes móviles ad hoc para zonas de recursos limitados. Investigaciones en Facultades de Ingeniería del NOA. Santiago del Estero: UNIVERSIDAD NACIONAL DE SANTIAGO DEL ESTERO. vol. n°. p105 - 106. issn 1853-7871.
7. HERRERA, S.; NAJAR RUIZ, P.; LEDESMA, E.; ROCABADO, S. (2012-a). Sistema de Información Móvil para Turismo Receptivo. Revista Gestao & Conhecimento. Pozos de Caldas: PUC Minas. vol. n°. p479 - 492. issn 1808-6594.
8. HERRERA, S.; ZUAIN, S.; GALLO, F.; AVILA, H. (2012-b). Emotion and Communicability in e-culture Applications. LECTURE NOTES IN COMPUTER SCIENCE. Heidelberg: Springer. vol.7546 n°. p15 - 24. issn 0302-9743.
9. HERRERA, S.; CLUSELLA, M.; MITRE, G.; SANTILLAN, A.; GARCIA, C. (2011-a). An interactive information system for e-culture. LECTURE NOTES IN COMPUTER SCIENCE. Heidelberg: Springer. vol.6616 n°. p30 - 43. issn 0302-9743.
10. HERRERA, S.; CLUSELLA, M.; LUNA, P.; MITRE, G.; SANTILLAN, A. SANTIAGUEÑIDAD SIGLO XXI. RETROPROSPECTIVACIÓN SISTÉMICA DE LA CULTURA SANTIAGUEÑA. (2011-b). Santiago del Estero: Lucrecia. pag.183. isbn 978-987-1754-81-6.
11. Ionic. (2016). Disponible en: <http://ionicframework.com/docs/guide/preface.html>. Fecha de acceso: 13/08/2016.
12. Najjar, P., Ledesma, E., Rocabado, S., Herrera, S., Palavecino, R. (2014). Eficiencia de aplicaciones móviles según su arquitectura. CACIC 2014. ISBN 978-987-3806-05-6. La Matanza, Buenos Aires.
13. PhoneGap 3. (2013). A guide to building cross-platform apps using the W3C standards-based Cordova/PhoneGap framework. Packt Publishing.
14. ROCABADO, S.; CADENA, C.; DIAZ, J. (2016-a). CARGADORES SOLARES PORTATILES PARA EL USO DE DISPOSITIVOS MOVILES EN ZONAS RURALES AISLADAS DEL NOA. AVERMA: INENCO. vol. n°. p - . issn 2314-1433.
15. ROCABADO, S.; SANCHEZ, E.; HERRERA, S.; CADENA, C. (2016-b). Eficiencia energética en dispositivos móviles para facilitar su uso en zonas rurales aisladas. Argentina. Concordia. Libro. Artículo Breve. Workshop. XVIII Workshop de Investigadores en Ciencias de la Computación. Red UNCI.
16. ROCABADO, S.; CADENA, C.; DIAZ, J. 2015. USO DE DISPOSITIVOS MÓVILES EN ZONAS RURALES AISLADAS ABASTECIDAS CON ENERGÍA SOLAR FOTOVOLTAICA. (2015). Argentina. JUNIN. Libro. Artículo Completo. Congreso. CACIC. Red de Universidades con Carreras en Informática RedUNCI.
17. ROCABADO, S.; HERRERA, S.; MORALES, M.; ESTELLES, C. (2013). M-LEARNING EN ZONAS DE RECURSOS LIMITADOS. Argentina. Santiago del Estero. Libro. Artículo Completo. Congreso. VIII Congreso de Tecnología en Educación y Educación en Tecnologías. Red de Universidades con Carreras de Informática (RedUNCI).
18. Sánchez Zuaín, S., Durán, E. (2016). Artículo Taxonomía de Requisitos para Aplicaciones Web Congreso CACIC. San Luis. En evaluación.
19. Siegler, GM (2008). Analyst: there is a great future in iPhone apps. Disponible en: <http://venturebeat.com/2008/06/11/analyst-theresa-great-future-in-iphone-apps/>. Consulta realizada el 08/07/2016.
20. Talukder, A., Ahmed, H., & Yavagal, R. (2010). Mobile Computing: Technology, Applications, and Service Creation. Second Edition. -: McGraw-Hill Professional.
21. Wargo, J. (2015). Apache Cordova 4 Programming. Addison-Wesley Professional

Un Método de Evaluación para Modelo de Gestión de Calidad en Empresas de SSI

Walter A. Lucero, Marina Gette, Carlos Salgado, Alberto Sánchez, Mario Peralta
Departamento de Informática Facultad de Ciencias Físico-Matemáticas y Naturales
Universidad Nacional de San Luis
Ejercito de los Andes 950 – C.P. 5700
San Luis - Argentina
{walteradrianlucero, marinagette}@gmail.com - {csalgado, alfanego, mperalta}@unsl.edu.ar

Resumen

Los modelos de calidad son referencias que las organizaciones utilizan para mejorar su gestión. Los modelos, a diferencia de las normas, no contienen requisitos que deben cumplir los sistemas de gestión de la calidad sino directrices para la mejora. Existen modelos de calidad orientados a la calidad total y la excelencia, modelos orientados a la mejora, modelos propios de determinados sectores e incluso modelos de calidad que desarrollan las propias organizaciones [1].

Por ello, se propone un modelo de Gestión de la Calidad orientado a empresas del Sector Software y Servicios Informáticos, basado en los 7 principios de la calidad de las organizaciones [1]: Enfoque al cliente, Liderazgo, Enfoque a procesos, Compromiso de las personas, Mejora, Toma de decisiones basada en la evidencia y Gestión de las relaciones.

El objetivo del modelo es brindar a las empresas de desarrollo de software de la región una herramienta que les permita posicionarse en un nivel altamente competitivo en el mercado actual, mediante la producción de software de calidad.

El objetivo del presente trabajo es evaluar, mediante un método englobado en las herramientas prospectivas como es el método Delphi, las características del

Modelo de Gestión de la Calidad orientado a empresas de SSI de la República Argentina.

Palabras claves: Método, Evaluación de Modelo de Gestión de Calidad en Empresas de SSI, Alfa de Crombach. Fiabilidad.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de incentivos.

Introducción

El método Delphi [3] se engloba dentro de los métodos de prospectiva, que estudian el futuro, en lo que se refiere a la evolución de los factores del entorno tecno-socio-económico y sus interacciones.

El primer estudio de Delphi fue realizado en 1950 por la Rand Corporation para la fuerza aérea de Estados Unidos, y se le dio el nombre de Proyecto Delphi. Su objetivo era la aplicación de la opinión de expertos en la selección de un sistema industrial norteamericano óptimo y la estimación del número de bombas requeridas para reducir la producción de municiones hasta un cierto monto.

Es un método de estructuración de un proceso de comunicación grupal que es efectivo a la hora de permitir a un grupo de individuos, como un todo, tratar un problema complejo [2].

En un artículo de Cabero e Infante, [9], se señala que: [...] el método Delphi es, posiblemente, uno de los más utilizados en los últimos tiempos por los investigadores para diferentes situaciones y problemáticas, que van desde la identificación de tópicos a investigar, especificar las preguntas de investigación, identificar una perspectiva teórica para la fundamentación de la investigación, seleccionar las variables de interés, identificar las relaciones causales entre factores, definir y validar los constructos, elaborar los instrumentos de análisis o recogida de información, o crear un lenguaje común para la discusión y gestión del conocimiento en un área científica. Es, por tanto, de verdadera utilidad para los investigadores de ciencias sociales en general, y los de educación y comunicación en particular.

El objetivo de los cuestionarios sucesivos es “disminuir el espacio intercuartil, esto es, cuánto se desvía la opinión del experto de la opinión del conjunto, precisando la mediana, de las respuestas obtenidas”.

La calidad de los resultados depende, sobre todo, del cuidado que se ponga en la elaboración del cuestionario y en la elección de los expertos consultados.

Este método se emplea bajo las siguientes condiciones:

- No existen datos históricos con los que trabajar.
- El impacto de los factores externos tiene más influencia en la evolución que el de los internos.
- Las consideraciones éticas y morales dominan sobre las económicas y tecnológicas en un proceso evolutivo.
- Cuando el problema no se presta para el uso de una técnica analítica precisa.
- Cuando se desea mantener la heterogeneidad de los participantes a fin de asegurar la validez de los resultados.
- Cuando el tema en estudio requiere de la participación de individuos expertos en distintas áreas del conocimiento.

A continuación, se muestran las fases con las respectivas etapas del método propuesto.

Características del Método para evaluar el Modelo de Gestión de Calidad para Empresas de SSI:

Fase I

Etapas 1. Estructura: El éxito del funcionamiento del método, radica en establecer en forma precisa la estructura del Modelo de Gestión de Calidad para empresas de SSI. En este punto se identifican cada uno de los componentes del Modelo a evaluar.

Etapas 2. Determinación del Grupo de Expertos: El objetivo principal en esta etapa es la elección del grupo de expertos.

La etapa es importante en cuanto que el término experto es “indeterminado”. Con independencia de sus títulos, función o nivel jerárquico, la elección del experto estará basada en la capacidad de encarar el futuro y el conocimiento profundo del tema.

Etapa 3. Periodicidad: La periodicidad está íntimamente relacionada con los recursos existentes en la organización y es decisión directa de la alta dirección.

Fase II

Etapa 4. Diseño y aplicación del cuestionario: En la etapa de diseño de un producto o servicio es de suma importancia saber que desea el “cliente”, de esta manera se logra la satisfacción final. En el diseño del cuestionario debe estar presente ésta premisa La información que se presenta a los expertos no es solo el punto de vista de la mayoría, sino que se presentan todas las opiniones indicando el grado de acuerdo que se ha obtenido.

Etapa 5. Aplicación del Alfa de Crombach [4]: El cuestionario elaborado debe ser ponderado para ello se utiliza la tabla de Likert [5].

Referencias de la tabla de Likert:

1. **TED:** Totalmente En Desacuerdo.
2. **PED:** Parcialmente En Desacuerdo.
3. **I:** Indiferente (no puede indicar acuerdo ni desacuerdo en forma precisa).
4. **PDA:** Parcialmente De Acuerdo.
5. **TDA:** Totalmente De Acuerdo.

Etapa 6. Evaluación de los resultados: En esta etapa se evalúan los resultados obtenidos de las etapas anteriores.

Etapa 7. Validación de la efectividad de los resultados por técnicas estadísticas:

Se utiliza una técnica estadística conocida como Alfa de Crombach [4].

Cálculo del Alfa de Crombach: El alfa de Crombach permite determinar la FIABILIDAD de la encuesta diseñada anteriormente, es decir, si realmente podemos confiar en la información que extraemos de la misma. De no ser fiable es necesario el rediseño de la encuesta.

Para comprobar la fiabilidad se toma una muestra piloto de 10 o 20 clientes y se calcula el coeficiente ALFA. Para el cálculo de este, es necesario que la encuesta esté diseñada bajo la escala de Likert [5].

$$\alpha = \frac{K}{K-1} \left[1 - \frac{\sum \sigma_i^2}{\sigma_w^2} \right]$$

Donde:

α = *coeficiente de Crombach*

σ_i^2 = *Varianza del ítem i*

σ_w^2 = *Varianza de la suma de todos los ítems*

Etapa 8. Acciones:

Todas las no conformidades encontradas vuelven al modelo para ser replanteadas.

Ventajas del Método:

- Permite obtener información de puntos de vista sobre el comportamiento del Modelo de Gestión de Calidad de SSI.
- El alfa de Cronbach es una herramienta estadística cuyo índice otorga la certeza para evaluar el grado en que los ítems de un instrumento están correlacionados.
- El horizonte de análisis puede ser variado.
- Ayuda a explorar de forma sistemática y objetiva problemas que

requieren la concurrencia y opinión cualificada.

Desventajas:

- Su elevado costo.
- Su tiempo de ejecución (desde el período de formulación hasta la obtención de los resultados finales), es intermedio.
- Es necesario tener conocimiento de análisis cuantitativo de datos para el tratamiento de la información.
- Se requiere de la utilización de un software estadístico para procesar la base de datos.

Líneas de Investigación, Desarrollo e Innovación

En los últimos años se evidenció un significativo incremento en la cantidad de empresas del sector de Software y Servicios Informáticos (SSI) en la República Argentina, según datos oficiales, actualizados por el Observatorio de Empleo y Dinámica Empresarial del Ministerio de Trabajo, el número de empresas creció un 132% en 10 años consecutivos [6].

Debido a este fenómeno, las empresas SSI se ven en la necesidad de diferenciarse de la competencia para captar nuevos clientes como así también mantener el crecimiento al ritmo de los avances tecnológicos y los requerimientos del mercado.

Muchas empresas del sector no se encuentran aplicando un modelo de gestión integrado que esté orientado a los clientes, el liderazgo, los procesos, los resultados y la mejora continua lo cual las aleja de la excelencia y, por ende, son poco productivas. Si bien un importante porcentaje de empresas de software, alrededor del 60% según los últimos

resultados publicados [6], poseen algún tipo de certificación, no cuentan con un modelo a seguir para su éxito a largo plazo.

Se propone en esta línea de investigación diseñar un Método basado en Técnicas Estadísticas como es el Alfa de Crombach, que valide el Modelo de Gestión de la Calidad para el sector Software y Servicios Informáticos.

Resultados y Objetivos

En base a todo lo expresado anteriormente, se definió un modelo formado por 6 componentes interrelacionados, los cuales pueden ser considerados por las empresas SSI de la República Argentina [7, 8] para la aplicación del Método propuesto.

Los ítems a valorar del Modelo son:

1. El Entorno de la Organización.
2. Estrategia y Política.
3. Gestión de los Recursos.
4. Gestión de los Procesos
5. Seguimiento, medición, análisis y revisión.
6. Mejora, innovación y aprendizaje

Se vio la necesidad de definir un método de evaluación para aplicar o evaluar el modelo de calidad previamente definido en la línea de investigación.

Entre los trabajos a futuro se planifica el desarrollo de herramientas de autoevaluación que sirvan de referencia para determinar el nivel de madurez actual de la empresa SSI y constituyan una base para la mejora continua.

Formación de Recursos Humanos

Bajo esta línea, en el grupo de investigación, se están desarrollando dos

tesis de maestría en Calidad de Software. En una de ellas se está definiendo/actualizando el modelo de calidad para las empresas del sector SSI basado en un enfoque de gestión por procesos orientado a los clientes, el liderazgo directivo, el personal, los resultados y la mejora continua.

La otra consiste en la definición de un Método de Evaluación para el Modelo de Gestión de Calidad en Empresas de SSI.

También se están llevando a cabo trabajos de grado con relación a la temática por alumnos de la carrera de Ingeniería Informática y Licenciatura en Ciencias de la Computación en la Universidad Nacional de San Luis.

Referencias

- [1] Asociación Española para la Calidad <http://www.aec.es/web/guest/centro-conocimiento/modelos-de-calidad>
- [2] Linstone, H. A., & Turoff, M. (1975). Introduction. In H. A. Linstone, & M. Turoff (Eds.). *The Delphi method: Techniques and applications* (pp. 3-3). Reading, MA: Addison-Wesley Publishing Company.
- [3] Landeta, J. (1999). *El método Delphi. Una técnica de previsión para la incertidumbre*. Barcelona: Ariel.
- [4] González Alonso, J. A., Pazmiño, M. Cálculo e interpretación del Alfa de Crombach para el caso de validación de la consistencia interna de un cuestionario, con dos posibles escalas tipo Likert. *Revista Publicando*, 2(1). 2015, 62-77. ISSN 1390-9304
- [5] Morales, P. (2013). El análisis factorial en la construcción e interpretación de tests, escalas i cuestionarios. Madrid. <http://web.upcomillas.es/personal/peter/investigacion/AnalisisFactorial.pdf>
- [6] Cámara de la Industria Argentina de Software, <http://www.cessi.org.ar/opssi>.
- [7] Marina C Gette, Alberto Sánchez, Carlos Salgado, Mario Peralta. “Modelo de Gestión de la Calidad orientado a empresas de Software y Servicios Informáticos (SSI) de la República Argentina.” En *Workshop de Investigadores en Ciencias de la Computación, WICC 2017*. Buenos Aires, Argentina.
- [8] Gette, M., Sanchez, A., Salgado, C., Peralta, M. 2017, Un Modelo de Gestión de la Calidad orientado a Empresas de Software y Servicios Informáticos (SSI) de la República Argentina. 5to Congreso Nacional de Ingeniería Informática - Sistemas de Información CoNaIISI 2017. ISSN 2347-0372
- [9] Cabero, J., Infante, A. Empleo del Método Delphi y su Empleo en la Investigación en Comunicación y Educación. *EDUTEC. Revista Electrónica de Tecnología Educativa*. ISSN 1135-9250. N° 48. 2014

INCLUSIÓN DE HACKING ÉTICO EN EL PROCESO DE TESTING DE SOFTWARE

Ariel Giannone, Sebastian Martins, Hernán Amatriain, Darío Rodríguez, Hernán Merlino

Laboratorio de Investigación y Desarrollo en Ingeniería de Explotación de Información
Grupo de Ingeniería de Explotación de Información y Grupo Investigación en Sistemas de Información

Departamento de Desarrollo Productivo y Tecnológico. Universidad Nacional de Lanús

29 de Septiembre 3901 (1826) Remedios de Escalada, Lanús. Argentina. Tel +54 11 5533 5600 Ext. 5194

giannoneariel@gmail.com, smartins089@gmail.com, hamatriain@gmail.com, dariorodriguez1977@gmail.com,
hmerlino@gmail.com

RESUMEN

Como ocurre con la mayoría de los avances tecnológicos, el crecimiento explosivo de Internet tiene un lado oscuro: los hackers. La escalada natural de amenazas ofensivas contra las medidas defensivas ha demostrado una y otra vez que no hay sistemas prácticos que se puedan construir que sean invulnerables a los ataques. Las organizaciones informatizadas se dieron cuenta de que una de las mejores formas de evaluar la amenaza de intrusión sería tener profesionales independientes de seguridad informática intentando entrar en sus sistemas. Estos "hackers éticos" emplean las mismas herramientas y técnicas que los intrusos, pero sin dañar el sistema de destino ni robar información. En su lugar, permiten evaluar la seguridad de los sistemas de destino e informar de a los propietarios sobre las vulnerabilidades encontradas junto con las instrucciones de cómo remediarlos. Este proceso debe ser planificado con antelación. Todos los aspectos técnicos, de gestión y estratégicos deben estar sumamente cuidados. Estas etapas deben realizarse en un marco de control, gestión y supervisión constante. Es allí donde apunta este proyecto, poder incluir de manera segura y metódica la fase de revisión por hacking ético dentro del proceso de Testing de software.

Palabras clave: Hacking Ético, Testeo de Software, Formación de Recursos Humanos.

CONTEXTO

El Proyecto articula líneas de investigación en el área de Seguridad de Espacios Virtuales de Trabajo del Laboratorio de Investigación y Desarrollo en Espacios Virtuales de Trabajo (LIDEVT UNLa) y Metodologías de Ingeniería de Software con radicación en el Departamento de Desarrollo Productivo y Tecnológico de la Universidad Nacional de Lanús. Las líneas de investigación del área cuentan con financiamiento de la Secretaría de Ciencia y Técnica de la misma Universidad.

FUNDAMENTACION

En general, las políticas de seguridad de la información o los controles por sí solos no garantizan la protección total de la información, ni de los sistemas de información, servicios o redes. Después de los controles que se han implementado, vulnerabilidades residuales probablemente permanezcan haciendo ineficaz la seguridad de la información y por lo tanto los incidentes son aún más posibles. Esto puede llegar a tener efectos negativos tanto directos e indirectos sobre las operaciones de negocio de una organización. Además, es inevitable que se produzcan nuevos casos de amenazas no identificadas previamente. Una preparación insuficiente por una organización para hacer frente a este tipo de incidentes hará cualquier respuesta menos efectiva, y aumentar así el grado de impacto comercial potencial adverso. [ISO/IEC 27035:2011]

En su búsqueda de una manera de abordar el problema, las organizaciones informatizadas se dieron cuenta de que una de las mejores formas de evaluar la amenaza de intrusión a sus intereses sería tener profesionales independientes de seguridad informática intentando entrar en sus sistemas. Este esquema es similar a tener auditores independientes entrando en una organización para verificar sus registros de contabilidad. En el caso de seguridad informática, estos "hackers éticos" emplean las mismas herramientas y técnicas que los intrusos, pero sin dañar el sistema de destino ni robar información. En su lugar, permiten evaluar la seguridad de los sistemas de destino e informar de a los propietarios sobre las vulnerabilidades encontradas junto con las instrucciones de cómo remediarlos.

En resumidas palabras, la evaluación de la seguridad de un sistema por parte de un hacker ético busca responder 3 preguntas básicas:

¿Qué puede ver un intruso en los sistemas atacados?
¿Qué puede hacer un intruso con esa información?

¿Hay alguien en el sistema atacado que se dé cuenta de los ataques o éxitos del intruso?

Este proceso debe ser planificado con antelación. Todos los aspectos técnicos, de gestión y estratégicos deben estar sumamente cuidados. La planificación es importante para cualquier todas las pruebas, ya sea desde un simple análisis de contraseña a una prueba de penetración completa en una aplicación web [Mayorga Jácome et al, 2015; Santos Castañeda, 2016; Onofa Calvopiña et al, 2016; López Vallejo, 2017]. El resguardo de datos debe garantizarse, de lo contrario la prueba puede volverse en contra si alguien afirma que nunca se autorizaron las pruebas. Por lo tanto, un alcance bien definido implica la siguiente información:

- Sistemas específicos para probar.
- Estimar los riesgos que están involucrados.
- Tiempo que llevara la prueba y evaluación del calendario general.
- Recoger y explorar el conocimiento de los sistemas que tenemos antes de la prueba.
- Entrega de informes específicos incluyendo informes de evaluación de la seguridad y un informe de nivel superior describiendo las vulnerabilidades generales que deben abordarse, junto con las medidas correctivas que se deben implementar.

Ahora bien el profesional de seguridad, al llevar a cabo un test de penetración como parte de su trabajo de hacking ético, necesita contar con ese tipo de lógica y tiene que aplicarla, más allá de utilizar las técnicas y herramientas open source [Comunidad Linux, 2014; OISSG, 2012; GNU, 2014], comerciales o privadas [Tenable Network Security, 2014], dado que necesita imitar un ataque de la mejor manera y con el máximo nivel posible [Coronel Suarez, 2016; Hurtado Sandoval et al, 2016; López Alvarez, 2016; López Vallejo, 2017]. Para eso, tendrá que emplear todos los recursos de inteligencia que tenga a su alcance, utilizar al extremo sus conocimientos, poder de deducción y análisis mediante el razonamiento y así determinar qué es lo mejor que puede intentar, cómo, dónde y con qué. Por ejemplo, saber si un pequeño dato, por más chico o insignificante que parezca, le será útil y cómo proseguir gracias a él. Continuamente se deberá enfrentar a etapas que le demanden la mayoría de estas aptitudes [Tori, 2008].

- Definir patrones de conducta y acción.
- Hacer relevamientos pasivos de información.
- Interpretar y generar código y cifrado de datos.

- Descubrir manualmente descuidos en el objetivo.
- Descubrir vulnerabilidades presentes de todo el escenario técnico.
- Proyectarse sobre la marcha en modo abstracto, táctica y estratégicamente.
- Ser exhaustivo, pero a la vez saber cuándo es el momento de recurrir a la distensión para no agotar la mente.

Ahora bien, estas etapas deben realizarse en un marco de control, gestión y supervisión constante la cual otorgue tranquilidad y seguridad tanto al profesional que se “coloca” en los pies del criminal como a la organización en su totalidad [Tamayo Veintimilla, 2016; Onofa Calvopiña et al, 2016; Paillacho Pozo et al, 2016]. Es allí donde apunta este trabajo, poder incluir de manera segura y metódica la fase de revisión por hacking ético dentro del proceso de Testing de software.

METODOLOGÍA DE DESARROLLO

Para construir el conocimiento asociado al presente proyecto de investigación, se seguirá un enfoque de investigación clásico [Riveros y Rosas, 1985] [Creswell, 2002] con énfasis en la producción de tecnologías [Sábato y Mackenzie, 1982]; identificando métodos, materiales y abordaje metodológico necesarios para desarrollar el proyecto:

Métodos

A continuación se definen los métodos que se llevarán a cabo en el presente trabajo. Ellos son:

Revisiones Sistemáticas

Las revisiones sistemáticas [Argimón, 2004] de artículos científicos siguen un método explícito para resumir la información sobre determinado tema o problema. Se diferencia de las revisiones narrativas en que provienen de una pregunta estructurada y de un protocolo previamente realizado.

Prototipado Evolutivo Experimental (Método de la Ingeniería)

El prototipado evolutivo experimental [Basili, 1993] consiste en desarrollar una solución inicial para un determinado problema, generando su refinamiento de manera evolutiva por prueba de aplicación de dicha solución a casos de estudio (problemáticas) de complejidad creciente. El proceso de refinamiento concluye al estabilizarse el prototipo en evolución.

Materiales

Para el desarrollo de los formalismos y procesos propuestos se utilizarán:

Formalismos de modelado conceptual usuales en la Ingeniería de Software [Rumbaugh et al., 1999][Jacobson et al., 2013] y en la Ingeniería del Conocimiento [García-Martínez y Britos, 2004].

Modelos de Proceso usuales en Ingeniería de Software [IEEE, 1997; ANSI/IEEE, 2007; Oktaba et al., 2007].

Metodología

Para alcanzar los Objetivos trazados se propone: (i) realizar una investigación documental exploratoria acerca de las técnicas de hacking ético para la evaluación de vulnerabilidades más utilizados, identificando casos de estudio y de validación, (ii) realizar una valoración de las técnicas de hacking ético para la evaluación de vulnerabilidades estudiadas en base a una comparación de cada una de sus características en los casos de estudio relevados, especificando detalladamente las vulnerabilidades explotadas por cada técnica y la afectación potencial a un sistema, elaborando un ranking de acuerdo a la peligrosidad de cada una, (iii) realizar un estudio detallado de las fases del proceso de testeo de sistemas para determinar las actividades donde deberían incorporarse las técnicas de hacking ético, (iv) proponer un modelo integrador de proceso de testeo que incluya el hacking ético para la evaluación de vulnerabilidades, definiendo fases, actividades y recomendando herramientas, (v) realizar pruebas de concepto en los casos de estudio y casos de validación identificados, que validen el método propuesto.

RESULTADOS ESPERADOS

El presente proyecto busca desarrollar e incorporar un método de hacking ético para la evaluación de vulnerabilidades dentro del procedimiento mismo de Testeo de un sistema. De esta manera, se aporta a los encargados de testing en sectores de Seguridad Informática de un grupo de actividades y herramientas que les brinde el soporte necesario para poder prevenir los problemas que en la actualidad son de creciente interés por las pérdidas económicas que conllevan.

FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo se encuentra formado por dos investigadores formados y tres investigadores en formación. En su marco se desarrollan dos Tesis de

Maestría y dos Especializaciones en Sistemas de Información.

REFERENCIAS

- ANSI/IEEE, (2007). Draft IEEE Standard for software and system test documentation. ANSI/IEEE Std P829-2007.
- Argimón J. (2004). Métodos de Investigación Clínica y Epidemiológica. Elsevier España, S.A. ISBN 9788481747096.
- Basili, V. (1993). The Experimental Paradigm in Software Engineering. En *Experimental Software Engineering Issues: Critical Assessment and Future Directions* (Ed. Rombach, H., Basili, V., Selby, R.). Lecture Notes in Computer Science, Vol. 706. ISBN 978-3-540-57092-9.
- Comunidad Linux. 2014 <<http://www.linux.org/>> Página Válida a 05/2017
- Coronel Suarez, I.A. (2016). Aplicar hackeo ético para detección de vulnerabilidades mediante herramientas Open Source en las aplicaciones web de una institución de educación superior. Tesis Maestría de la Escuela Superior Politécnica del Litoral. Repositorio Digital URI: <<http://www.dspace.espol.edu.ec/xmlui/handle/123456789/37397>>
- Creswell, J. (2002). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Prentice Hall. ISBN 10: 01-3613-550-1.
- Evans, Bob (2001). The Sorry State of Software. *InformationWeek* 112.
- García Martínez, R., Britos, P. (2004). Ingeniería de Sistemas Expertos. Editorial Nueva Librería. ISBN 987-1104-15-4.
- GNU (2014) Operating System Sponsored by the Free Software Foundation. 2014/05/15 <<http://www.gnu.org/>> Página Válida a 05/2017
- Hurtado Sandoval, M.E., Mendaño Mendaño, L.A. (2016). Implementación de técnicas de hacking ético para el descubrimiento y evaluación de vulnerabilidades de la red de una cartera de Estado. Tesis Grado. Escuela Politécnica Nacional de Ecuador. Repositorio Digital URI: <<http://bibdigital.epn.edu.ec/handle/15000/16836>>
- ISO 27035:2011 (2010). Information technology – Security techniques – Information security incident management. [Online]. <http://www.iso.org/iso/catalogue_detail?csnumber=44379>. Página Válida a 05/2017.
- IEEE, (1997). IEEE Standard for Developing Software Life Cycle Processes. IEEE Std 1074-1997 (Revision of IEEE Std 1074-1995; Replaces IEEE Std 1074.1-1995)
- Jacobson, I., Ng, P. W., McMahan, P. E., & Jaramillo, C. M. Z. (2013). La esencia de la ingeniería de software: El núcleo de Semat. *Revista Latinoamericana de Ingeniería de Software*, 1(3), 71-78.
- López Alvarez, D.M. (2015). Hacking ético para detección de vulnerabilidades de una empresa del sector de telecomunicaciones. Tesis Maestría de la Escuela Superior Politécnica del Litoral. Repositorio Digital URI: <<http://www.dspace.espol.edu.ec/xmlui/handle/123456789/36478>>
- López Vallejo, M.R. (2017) Hacking ético. Vulnerabilidad de Sistemas Operativos en el acceso por contraseñas. *Revista Publicando*, 4 No 10. (1). 2017, 31-51. ISSN

- 1390-9304.
<http://rmlconsultores.com/revista/index.php/crv/article/view/407/pdf_259> Página Válida a 05/2017
- Mayorga Jácome, T., Quisaguano Belduma, F.J. (2015). Implementación de hacking ético para el análisis de vulnerabilidades, métodos de prevención y protección aplicados a la infraestructura de red de la empresa Construlec Cía. Ltda. en Quito Ecuador. Editorial: Quito: Universidad Israel. Repositorio Digital URI: <<http://repositorio.uisrael.edu.ec/handle/47000/647>>
- OISSG (2012) Open Information Systems Security Group. 2003 – 2012 <<http://www.oissg.org/>> Página Válida a 05/2017
- Oktaba, H., Garcia, F., Piattini, M., Ruiz, F., Pino, F., Alquicira, C. (2007). Software Process Improvement: The Competisofit Project. IEEE Computer, 40(10): 21-28. ISSN 0018-9162.
- Onofa Calvopiña, F.O., Pilatuña Chica, I. (2016). Análisis y evaluación de riesgos y vulnerabilidades del nuevo portal web de la Escuela Politécnica Nacional, utilizando metodologías de hackeo ético. Tesis Grado. Escuela Politécnica Nacional de Ecuador. Repositorio Digital URI: <<http://bibdigital.epn.edu.ec/handle/15000/16740>>
- OWASP Top 10 - 2013 (2013). [Online]. <http://www.owasp.org>. Página Válida a 05/2017.
- Paillacho Pozo, P.A. (2016). Análisis, diseño y pruebas de hacking ético sobre la infraestructura institucional de la Superintendencia de Control del Poder de Mercado. Tesis Grado. Universidad Politécnica Salesiana de Ecuador. Repositorio Digital URI: <<http://dspace.ups.edu.ec/handle/123456789/13425>>
- Palmer, Charles (2001). Ethical hacking, IBM Systems Journal, Vol. 40, N°3
- Raymond E (1991). The New Hacker's Dictionary, MIT Press, Cambridge, MA
- Riveros, H. y Rosas, L. (1985). El Método Científico Aplicado a las Ciencias Experimentales. Editorial Trillas. México. ISBN 96-8243-893-4.
- Rumbaugh, J., Jacobson, I., Booch, G. (1999). The Unified Modeling Language, Reference Manual. Addison Wesley, ISBN-10: 02-0130-998-X.
- Sabato J, Mackenzie M. (1982). La Producción de Tecnología: Autónoma o Transnacional. Instituto Latinoamericano de Estudios Transnacionales - Technology & Engineering. ISBN 9789684293489.
- Santos Castañeda, D.M. (2016). Análisis y diagnóstico de vulnerabilidades informáticas en la red de datos de la empresa YOUPHONE Cía. Ltda. Utilizando Hacking Ético. Tesis Grado. Institucional de la Universidad de las Fuerzas Armadas ESPE. Repositorio Digital URI: <<http://repositorio.espe.edu.ec/handle/21000/12142>>
- Schneier, Bruce. (2000). Secrets & Lies: Digital Security in a Networked World. John Wiley & Sons.
- Sheoran, Pankaj & Singh, Sukhwinder (2014). Applications of Ethical Hacking, International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463 Vol. 3 Issue 5, May-2014, pp: (112-114), Impact Factor: 1.252, Available online at: www.erpublications.com Page | 112
- Software Engineering Institute (2014) – Carnegie Mellon University. 2014 <http://www.cert.org/tech_tips/malicious_code_mitigation.html> Página Válida a 05/2017
- Tamayo Veintimilla, O.A. (2016). Desarrollo de una guía técnica estándar para aplicar herramientas de Ethical Hacking en redes de datos, dirigido a PYMES. Tesis Grado de Pontificia Universidad Católica de Ecuador. Facultad de Ingeniería. Escuela de Sistemas. Repositorio Digital URI: <<http://repositorio.puce.edu.ec/handle/22000/12612>>
- Tenable Network Security (2014), proveedora de la herramienta Nessus, 2014 <<http://www.tenable.com/products/nessus?gclid=CK2xwJGavr4CFScHwwod9VMAZA>> Página Válida a 05/2017
- Tori C. (2008). Hacking Ético (1ra Ed). Buenos Aires: Mastroianni.
- Zimmerman, Christine. (2001). Race to Deploy May Magnify Software Bugs. InternetWeek 13.

ACCESIBILIDAD WEB, APORTANDO A LA INCLUSIÓN.

Sonia I. Mariño, Pedro L. Alfonzo, Cintia Galain, Josue Maidana, Romina Alderete
Departamento de Informática. Facultad de Ciencias Exactas y Naturales y
Agrimensura.

9 de Julio 1449. CP: 3400. Corrientes. Argentina.

Universidad Nacional del Nordeste.

simarinio@yahoo.com, plalfonzo@hotmail.com

RESUMEN

La Accesibilidad Web es un aspecto tecnológico de connotación social. En el trabajo se presentan avances en el fortalecimiento de la formación de recursos humanos en temas de Accesibilidad Web, con miras a la mejora sustantiva de e-soluciones para los ciudadanos dado que es un aspecto de la calidad de los sistemas informáticos.

Palabras clave: Accesibilidad Web, métodos y herramientas, formación de recursos humanos, transferencia de conocimientos.

CONTEXTO

En el marco de un proyecto de I+D acreditado por la Secretaria General de Ciencia y Técnica (UNNE), se indaga y aplican métodos y herramientas para evaluar la accesibilidad web en dispositivos móviles y considerado como un aspecto de la calidad de la Ingeniería del Software.

1. INTRODUCCIÓN

La evolución de las tecnologías web y móviles transforma actividades personales y profesionales. Además, son cada vez más sofisticados los dispositivos y el crecimiento tecnológico permite ejecutar aplicaciones más complejas [1]. Se coincide con [2] en que el uso masivo de dispositivos móviles crea un nuevo mercado para desarrolladores de software y genera nuevos desafíos para mejorar la vida cotidiana de las personas.

Por lo expuesto, es relevante asegurar la Accesibilidad Web (AW) en los productos software dado que se utilizan por usuarios con diferentes capacidades.

Para [3], en la Ingeniería del Software (IS) existen tres elementos claves: i) los métodos,

ii) las herramientas y iii) los procedimientos. Estos elementos facilitan el control del proceso de construcción de software y brindan a los desarrolladores las bases de la calidad de una forma productiva.

En este sentido, desde la Ingeniería del Software es posible determinar la calidad de los productos software en proceso de elaboración, siendo la Accesibilidad Web una medida aplicable desde etapas tempranas de desarrollo y tratada como un requerimiento no funcional [4].

La Accesibilidad Web referencia el acceso universal a este servicio de Internet, independientemente del tipo de hardware, software, infraestructura de red, idioma, cultura, localización geográfica y capacidades de los usuarios [5, 6].

La AW, desde la IS, aborda cómo se debe codificar y presentar la información cuando se diseña un sitio para lograr que las personas con o sin alguna discapacidad puedan percibir, entender, navegar e interactuar de forma efectiva con la Web, así como también crear y aportar contenido [7].

En el marco del proyecto “TI en los Sistemas de Información: modelos, métodos y herramientas” se avanza en la indagación de métodos y herramientas y su aplicación con miras a aportar a la inclusión de los ciudadanos en el uso de herramientas informáticas en este siglo, y contribuir a que las Tecnologías de la Información y Comunicación (TIC) estén al servicio de la comunidad para mejorar su calidad de vida.

Para el equipo es fundamental el estudio teórico y la definición de procedimientos orientados a aplicar la AW en los productos tecnológicos.

Cabe aclarar que la iniciativa de estudio de AW, se trata en otras universidades como se mencionan en [1, 2, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17].

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Con el objetivo de ampliar conocimientos teóricos y desarrollos empíricos en la temática se fomenta la construcción de aplicaciones móviles que respondan a estándares internacionales de accesibilidad como los establecidos por la WCAG 2.0 se:

- Indagación en métodos para el diseño de soluciones web y móviles accesibles.
- Estudio de estándares de medición de la AW.
- Relevamiento de herramientas para la medición de accesibilidad web y móvil.
- Validación de herramientas en distintas plataformas.

3. RESULTADOS Y OBJETIVOS

Como se expresó previamente [18] y siguiendo los lineamientos de la RedUNCI [19], en el proyecto se aborda la Accesibilidad Web en el desarrollo en el grado y posgrado, con la finalidad de contribuir desde la Universidad con la formación de recursos humanos que se insertan en la Industria del Software. Es así como se logra:

- la elaboración de dos planes de trabajo para becas de pregrado orientadas a la revisión y profundización en métodos y herramientas de Accesibilidad Web y la introducción de su estudio desde etapas tempranas del ciclo de vida de las aplicaciones móviles.
- la aprobación de proyectos de tesis de posgrado vinculadas a la AW en dominios de la Educación.
- la aprobación del plan de trabajo de un Proyecto Final de Carrera [20], modalidad trabajo en equipo.

En referencia a los avances tecnológicos se mencionan:

- identificación y determinación de instrumentos para evaluar la AW, como sustento de nuevas propuestas.
- estudio, examen y aplicación de métodos para el tratamiento de la accesibilidad web basados en las pautas descritas en [21] y adaptadas de [22].
- elección, análisis estudio y aplicación de herramientas informáticas para la medición de accesibilidad móvil, entre las que se mencionan: Test de Accesibilidad de Google [23].
- utilización de diversos dispositivos que responden a distintas configuraciones para evaluar el nivel de accesibilidad de las aplicaciones.
- aplicación de las pautas WCAG 2.0 [22], desde las etapas iniciales del desarrollo de las Apps, particularmente se aplica en el diseño de un producto destinado al turismo.

4. FORMACIÓN DE RECURSOS HUMANOS

La universidad se desempeña responsable socialmente, así contribuye en la formación de recursos humanos de acuerdo a las exigencias de las empresas y gobiernos, generando y potenciando los vínculos entre Universidad-Empresa-Estado.

Los desarrollos tecnológicos diseñados y construidos con recursos humanos formados y en formación aportan a concretar la meta. En el año 2017 se fortalece la formación de recursos humanos desde el grado incorporando becarios [25, 26] y en el posgrado con trabajos de investigación aplicada.

Como líneas futuras de trabajo se menciona desarrollar asesorías con la finalidad de aportar, desde la producción de software en la conformación de una sociedad inclusiva centrada en los sujetos usuarios de la tecnología.

5. BIBLIOGRAFIA

- [1] P. Thomas, F. Cristina, S. Dapoto y P. Pesado, “Desarrollo de Aplicaciones Móviles 3D.”, presentado en XIX Workshop de Investigadores en Ciencias de la Computación, 2017.
- [2] J. Fernández Sosa, A. Cuitiño, P. Thomas, L. Delia, G. Cáseres, L. Corbalán, y P. Pesado, “Informática UNLP, la App de la Facultad de Informática.”, presentado en XXIII Congreso Argentino de Ciencias de la Computación, 2017.
- [3] R. S. Pressman, “Ingeniería de Software: Un Enfoque Práctico.”, Madrid: Pearson Education, S.A.
- [4] S. I. Mariño, M. V. Godoy, P. Alfonzo, J. Acevedo, L. Gómez Solís, A. Fernández Vázquez, “Accesibilidad en la definición de requerimientos no funcionales. Revisión de herramientas.”, *Multiciencias*, 12(3), pp. 305-312, 2012.
- [5] Consorcio World Wide Web (W3C). [Online]. Disponible: <http://www.w3c.es/>
- [6] Oficina Española. “Word Wide Web - Guía Breve de Accesibilidad Web”, [Online]. Disponible: <http://www.w3c.es/divulgacion/guiasbreves/accesibilidad>
- [7] S. Mariño, P. Alfonzo, “Evaluación de la accesibilidad web. Una mirada para asegurar la formación en la temática”. *Campus Virtuales: Revista científica iberoamericana de tecnología educativa*, 6(2), pp. 21-30, 2017.
- [8] J. S. Filippi, H. D. Perez, S. Aguirre. y Bertone, R. (2017). “ReadMe. Complemento de Aprendizaje Móvil.”, presentado en XXIII Congreso Argentino de Ciencias de la Computación, 2017.
- [9] C. Challiol, A. Lliteras, y S. E. Gordillo, “Diseño de Aplicaciones Móviles basadas en Posicionamiento: un Framework Conceptual.” presentado en XXIII Congreso Argentino de Ciencias de la Computación, 2017.
- [10] F. Cristina, S. Dapoto, P. Thomas y P. Pesado, “Evaluación de performance de engine 3D para dispositivos móviles.”, presentado en XXIII Congreso Argentino de Ciencias de la Computación, 2017.
- [11] M. Bustos, N. Perez, y M. Berón “Tecnología Mobile Aplicada a las Instituciones Educativas.”, presentado en *XIX Workshop de Investigadores en Ciencias de la Computación*, 2017.
- [12] C. Chayle, C. M. Herrera, M. A. Barrera, A. C. Pauletto y S. D. Blanco, 2017. “Evaluación de la Accesibilidad Web. Caso de Estudio: Sitios Web de la UNCA.”, presentado *XIX Workshop de Investigadores en Ciencias de la Computación*, 2017.
- [13] V. Castro, C. Ortiz, V. Chapetto, C. Balleto, y B. Rossi, “¿Las Redes Sociales Cumplen con los Criterios de Accesibilidad?.”, presentado en XIX Workshop de Investigadores en Ciencias de la Computación, 2017.
- [14] T. Barrios, M. Marín, y N. Torrente, “El Uso de la Tecnología para la Inclusión de los Disminuidos Visuales en las Aulas”, presentado en XIX Workshop de Investigadores en Ciencias de la Computación, 2017.
- [15] C. J. Reyes, M.L. Massé Palermo, C. Espinoza, C. Vargas, J. Ramírez, y J.E. Trenti, “Dispositivos Móviles como Soporte para el Aprendizaje Colaborativo de Programación en el Nivel Universitario Inicial (resultados).”, presentado en XIX Workshop de Investigadores en Ciencias de la Computación, 2017.

- [16] R. Rodríguez, P. Vera, R. Martínez, F. Parra Beltrán y J. Alcidor, “Análisis e Implementación de Nuevas Tecnologías para la Web Móvil.”, presentado en XIX Workshop de Investigadores en Ciencias de la Computación, 2017.
- [17] P. Thomas, L. Delia, L. Corbalan, G. Cáseres, N. Galdamez, A. Cuitiño, J. Sosa, y P. Pesado, “Análisis de Enfoques de Aplicaciones para Dispositivos Móviles.”, presentado en XIX Workshop de Investigadores en Ciencias de la Computación, 2017.
- [18] S. I. Mariño, M. V. Godoy, P. Alfonso, “Avances en torno a la formación en accesibilidad web”, XIX Workshop de Investigadores en Ciencias de la Computación, 2017, p. 687-690.
- [19] RedUNCI, Documento de Recomendaciones Curriculares de la RedUNCI2014-2015. [Online] Disponible: <http://redunci.info.unlp.edu.ar>,
- [20] C. Galain, J. Maidana, “Epuen 2.0. Una App accesible para el turismo local”, Proyecto Final de Carrera. Carrera Licenciatura en Sistemas de Información, FaCENA, UNNE, Corrientes, 2017.
- [21] Ilunion, “Metodología para evaluar la accesibilidad de aplicaciones móviles”, 2015, [Online]. Disponible: http://www.amovil.es/sites/default/files/metodologia_para_evaluar_la_accesibilidad_de_aplicaciones_nativas.pdf
- [22] Web Content Accessibility Guidelines (WCAG) 2.0. [Online] Disponible: <http://www.w3.org/TR/WCAG20/>
- [23] Test de Accesibilidad. [Online] Disponible: <https://play.google.com/store/apps/details?hl=es&id=com.google.android.apps.accessibility.auditor>
- [24] Google TalkBack. [Online] Disponible: <https://play.google.com/store/apps/details?id=com.google.android.marvin.talkback&hl=es>
- [25] J. Maidana, “TIC y GC. Métodos y herramientas para producción de una App de alcance regional.” Propuesta de beca para EVC-CIN.
- [26] C. Galain, “TI y GC. Una propuesta de aplicación móvil para la difusión del turismo.” Beca de Estímulo a la Investigación Científica – FACENA, UNNE.

DISEÑO DE SOFTWARE BASADO EN BPM PARA MEJORAR LA USABILIDAD DE LAS APLICACIONES

Company, Ana M.; Dapozo, Gladys N.; Irrazabal, Emanuel
Departamento de Informática. Facultad de Ciencias Exactas y Naturales y
Agrimensura. Universidad Nacional del Nordeste
anamacom@gmail.com; {gndapozo, eirrazabal}@exa.unne.edu.ar

RESUMEN

Debido a la necesidad de construir software con mayor cantidad de funcionalidades, donde esas funcionalidades se refieren a procesos críticos, que muchas veces complejizan el uso de los sistemas, en este proyecto se propone la elaboración de un marco de trabajo para el diseño de software basado en procesos de negocio, que permita mejorar la usabilidad de los sistemas, particularmente en el ámbito de un organismo del estado provincial.

Palabras clave: Modelado de Procesos de negocio. BPM. BPMN. Usabilidad.

CONTEXTO

Esta propuesta es parte del plan de Trabajo Final de la Maestría en Tecnologías de la Información (UNNE) que se desarrolla en el marco del proyecto 17F018 “Metodologías y herramientas emergentes para contribuir con la calidad del software”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE).

1. INTRODUCCIÓN

A medida que se requiere la construcción de software con mayor cantidad de funcionalidades, donde esas funcionalidades se refieren a procesos críticos, es importante comenzar a pensar en nuevas formas de diseñar aplicaciones que permitan la gestión efectiva de los diferentes procesos involucrados en el negocio del sistema, mejorando la facilidad de uso de las aplicaciones. Los problemas de usabilidad de los sistemas de información han sido ampliamente estudiados en los últimos años,

intentando encontrar la mejor manera de hacer que los sistemas de información sean más fáciles y efectivos de usar. El uso de modelos de procesos de negocio es una de las áreas de investigación que podría generar nuevas ideas para mejorar la usabilidad de los sistemas de información [1].

La visión de proceso de negocio resulta especialmente adecuada para las organizaciones estructuradas con objetivos bien definidos, como los organismos del estado. En este caso, se propone orientar la solución al Instituto de Vivienda de Corrientes (INVICO), un ente autárquico de la provincia de Corrientes, cuya misión es satisfacer las necesidades de hábitat y calidad de vida de todos los ciudadanos construyendo espacios adecuados para su desarrollo.

El instituto gestiona:

- Inscripciones de postulantes a acceder a una solución habitacional
- Sorteo de viviendas
- Créditos para adjudicatarios de una solución habitacional INVICO, incluyendo el proceso de recupero de cuotas
- Obras / Certificación de obras
- Proveedores / Contratistas
- Recursos Humanos

Maneja grandes cantidades de información, que son de vital importancia para el funcionamiento del instituto.

El organismo cuenta con el Departamento Sistemas, compuesto por las áreas de Desarrollo de Software, Mesa de Ayuda, Infraestructura y Bases de Datos, y Comunicación Visual.

El área de Desarrollo de Software se dedica tanto a mantenimiento de sistemas existentes

como a desarrollo de nuevos sistemas, que forman parte de un plan integral de actualización de tecnología (migración de sistemas existentes) para lograr la integración completa de todos los sistemas INVICO. Estas actualizaciones de tecnologías se deben a que todavía quedan algunos sistemas antiguos los cuáles no están conectados entre sí, además de que algunos sistemas se han vuelto obsoletos.

El Departamento de Sistemas busca la mejora continua de las aplicaciones software, a través de la aplicación de las últimas tecnologías que permitan mejorar las funcionalidades como así también la interacción del usuario a través de interfaces intuitivas usables que mejoren la experiencia del usuario con los sistemas.

En el día a día, el equipo de desarrollo, trabaja sobre un proyecto de software nuevo a la vez, y además se ocupa de solucionar requerimientos de los sistemas existentes, solicitados por las diferentes áreas del Instituto.

Se trata de optimizar el proceso de desarrollo para reducir los tiempos de las entregas y lograr mejoras en la distribución del trabajo del equipo.

Para lograr este objetivo se propone la elaboración de un marco de trabajo, basado en el modelado de los procesos de negocio, que oriente el diseño de las aplicaciones para mejorar la usabilidad de los sistemas, que permita disminuir el tiempo de aprendizaje de utilización por parte de los usuarios.

A. Trabajos Relacionados

En [1] los autores proponen una solución innovadora para la gestión y visualización de la ejecución de procesos de negocio. A diferencia de los sistemas de información tradicionales donde la ejecución de procesos de negocios se lleva a cabo usando controles de interfaz (ventanas, pantallas, menús), se propone un enfoque para administrar la ejecución de procesos de negocios directamente usando diagramas gráficos que describen procesos de negocios. El usuario

puede iniciar las actividades de ejecución del proceso directamente haciendo clic en el elemento deseado del diagrama del proceso de negocio. De esta forma, se pueden iniciar todas las actividades relevantes: iniciar la ejecución del proceso, elegir la ruta de ejecución, abrir objetos/documentos de datos, finalizar la ejecución del proceso, etc. Las actividades ejecutadas se pueden representar visualmente en los diagramas de procesos de negocio, por ejemplo, coloreando los pasos del proceso de negocio ya ejecutados. Además, los autores aseguran que el mecanismo de ejecución del proceso comercial propuesto permite mejorar significativamente la usabilidad de los sistemas de información.

B. BPM (Business Process Management)

BPM es un conjunto de métodos, herramientas y tecnologías utilizados para diseñar, representar, analizar y controlar procesos de negocio. Promueve la colaboración entre personas de negocio y tecnólogos para fomentar procesos de negocio efectivos, ágiles y transparentes [2].

Para el modelado de procesos de negocio se utilizan diagramas gráficos que se organizan jerárquicamente. Cada diagrama en la jerarquía describe un objeto o actividad de un nivel superior. Los diagramas se componen de símbolos gráficos que representan las actividades necesarias para la ejecución de los procesos de negocio, como los objetos de datos, y el flujo de control y de datos, además de los diferentes momentos en el tiempo. En [1] se plantea la posibilidad de trabajar con elementos gráficos de diagramas directamente, para lo que se cuestionan la identificación de qué requisitos debe cumplir un modelo de proceso de negocios para la ejecución directa de los procesos de información.

El uso de conceptos de modelado de procesos permite a los diseñadores especificar requisitos de proceso en términos de interacciones en la web, representadas por agentes humanos. El uso de servicios web

permite a los diseñadores modelar los requisitos de distribución de procesos derivados de las limitaciones de la organización, las oportunidades de diseño o los sistemas heredados existentes [3].

C. BPMN (*Business Process Modeling Notation*)

BPMN es un lenguaje para el modelado de procesos, que sigue una notación gráfica estándar gestionado por la OMG (Object Management Group), que permite la creación de diagramas de negocio mediante una técnica para graficar flujos de trabajo.

Tiene como objetivo proporcionar un lenguaje común para los diferentes usuarios involucrados en el negocio (analistas que crean y refinan los procesos, desarrolladores que implementan esos procesos, y directores que monitorizan y gestionan las actividades) [2].

Las grandes organizaciones modelan sus procesos en lenguajes como BPMN e incluyen programas para la mejora de procesos, aunque solo algunas usan sistemas BPM para ejecutar automáticamente sus procesos operativos. En la actualidad, es común encontrar en las organizaciones un enfoque orientado a "pensar en el proceso" [4].

D. Usabilidad

De las definiciones de Redish y Dumas, Krug y Nielsen, se desprende que la usabilidad tiene que ver con centrarse en el usuario que utiliza el software, proporcionar software adecuado para cualquier usuario independientemente de las habilidades, que le permita realizar sus tareas fácilmente, de manera rápida, eficiente y productiva.

Entonces, al desarrollar un producto, tenemos que pensar en la diversidad de usuarios posibles y con diferentes habilidades, conocimientos, y características. Diseñar para usuarios experimentados es bastante difícil, pero diseñar para usuarios no calificados es un desafío mucho mayor [5].

Para lograr interfaces amigables con el

usuario es necesario comprender los problemas de los usuarios frustrados, y de los que aún no se acercan a la tecnología. También es importante una mejor comprensión de las capacidades y limitaciones de las interfaces de usuario actuales.

En [5] se presenta el concepto de Usabilidad Universal, que tiene que ver con cómo pueden los servicios de información y comunicaciones ser utilizables para todos los ciudadanos. La usabilidad universal, podría plantearse como un concepto más transversal que la mera usabilidad, en el que se tratan conjuntamente aspectos de accesibilidad, user-friendly y usabilidad.

Las interfaces de usuario de computadoras incrementaron su importancia con el aumento del número de usuarios y aplicaciones, y una alta usabilidad es deseable, la que no aparece por arte de magia, sino que, para garantizar la facilidad de uso de los productos informáticos interactivos, se debe incluir activamente a las inquietudes de usabilidad en el proceso de desarrollo de software [6].

En [6] se presenta un modelo de usabilidad, versión modificada y ampliada de las "Reglas de oro" de Gould y Lewis, enfocado desde el principio en los usuarios, y su participación en el diseño y coordinación en las diferentes partes de la interfaz de usuario. Los elementos más básicos en el modelo de ingeniería de usabilidad son las pruebas de usuario empíricas y la creación de prototipos, combinado con un diseño iterativo.

Es importante tener en cuenta no sólo que un diseño de interfaz cumpla con las necesidades actuales, sino también si entra en conflicto con las habilidades que los usuarios han adquirido de las interfaces anteriores y si parece lo suficientemente flexible como para ser extendida para las interfaces futuras.

Entre las cinco características principales de usabilidad se encuentran [6]:

- Facilidad de aprendizaje
- Eficiencia en el uso, una vez que el sistema se ha aprendido

- Capacidad de los usuarios poco frecuentes para volver al sistema sin tener que aprenderlo todo
- Frecuencia y gravedad de los errores del usuario
- Satisfacción del usuario (subjetiva)

Además, existen nueve heurísticas de usabilidad [6]:

- Utilice el diálogo simple y natural
- Hable el idioma del usuario
- Reducir al mínimo la carga de memoria de usuario
- Se consistente
- Suministre realimentación
- Proporcionar salidas claramente marcadas
- Proporcione atajos
- Proporcionar buenos mensajes de error
- Evitar errores

En este proyecto, el foco se encuentra en la unión de ambos conceptos, procesos de negocio y usabilidad. Así, en [7] se presenta el estado del arte que estudia la combinación entre el modelado de procesos de negocio y el modelado de interfaces de usuario. Allí se concluye que es posible guiar el proceso de construcción de una solución y tener en cuenta tanto la ejecución del proceso como la interfaz de usuario en una etapa lo suficientemente temprana. Y que este enfoque mejora el impacto en la solución de trabajo.

Las investigaciones en el diseño de interfaz de usuario basado en modelos han generado diferentes tipos de lenguajes y estrategias específicas. Entre las más utilizadas, se encuentran las estrategias que se centran en el conocimiento del usuario y en el conocimiento que tenga el diseñador sobre el usuario, sobre sus objetivos y sobre las tareas que necesitan realizar para cumplir estos objetivos [8].

A su vez, para ello es necesario tener en cuenta la usabilidad del producto software resultante. En este sentido, la norma ISO/IEC

25010 [9] identifica seis características relacionadas con la usabilidad entre las que se pueden destacar la operatividad y el aprendizaje. Por un lado, la operatividad que busca medir la capacidad del producto software para ser controlado con facilidad. Por otro lado, el aprendizaje que consiste en la capacidad del producto software resultante para que los usuarios aprendan a utilizarlo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este proyecto se orienta a la elaboración de un marco de trabajo para el diseño de software basado en procesos de negocio, que permita mejorar la usabilidad de los sistemas en el ámbito de un organismo del estado provincial, tomando como caso de estudio un área de la gerencia de Programas Autogestivos del INVICO. Para lo que será necesario:

- Estudiar en profundidad los conceptos, técnicas y herramientas vinculadas con el modelado de procesos de negocio; y con métodos y técnicas para la mejora de la usabilidad desde el diseño de las interfaces de los sistemas de información.
- Identificar los procesos de negocio relevantes del área caso de estudio.
- Modelar los mismos utilizando el enfoque centrado en procesos BPM, mediante el estándar BPMN.
- Definir un marco de trabajo que basado en el modelado de los procesos de negocio identificados y en las técnicas de usabilidad, oriente el diseño de las aplicaciones.
- Validar y evaluar la propuesta mediante el desarrollo de una aplicación para el área caso de estudio.

3. RESULTADOS

OBTENIDOS/ESPERADOS

Se espera que el marco de trabajo resultante de este proyecto, se implemente en el Departamento Sistemas del INVICO para el desarrollo y/o mantenimiento del software de

gestión del instituto.

Como valor agregado de esta propuesta, el modelado de los procesos de negocios permitirá alinear los sistemas informáticos a los objetivos estratégicos de la organización.

La profundización en técnicas y métodos para mejorar la usabilidad del software, permitirá además mejorar el tiempo entre el análisis y el diseño, haciendo más eficiente el trabajo del equipo de desarrollo.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea participan dos investigadores formados y una tesista de posgrado que con esta propuesta espera lograr la obtención del título de Magister en Tecnologías de la Información, proyecto interinstitucional entre la Universidad Nacional de Misiones y la Universidad Nacional del Nordeste que se dicta en ambas universidades con un único plan de estudio.

5. BIBLIOGRAFÍA

- [1] J. Bicevskisa, Z. Bicevskab, Business Process Models and Information System Usability. University of Latvia, Riga, 2015.
- [2] P. F. Beltrán Naranjo, J.G. Yandún Montenegro, Análisis comparativo de herramientas BPM (Business Process Manager) orientadas por sector o línea de negocio. Trabajo de Titulación presentado de conformidad con los requisitos establecidos para optar por el título de Magister en Gerencia de Sistemas y Tecnologías de la Información. 2014.
- [3] M. Brambilla, S. Ceri, P. Fraternali, Process Modeling in Web Applications. Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy Ioana Manolescu INRIA Futurs - LRI, PCRI, France
- [4] W. M. P. Van Der Aalst, M. La Rosa, F. M. Santoro, Business Process Management. Don't Forget to Improve the Process!
- [5] B. Shneiderman, Universal Usability. Pushing human-computer interaction research to empower every citizen. Communications of the ACM. May 2000/vol. 43, no. 5
- [6] J. Nielsen, The Usability Engineering Life Cycle. Bellcore. 1992.
- [7] H. Trætterberg, J. Krogstie, Enhancing the usability of bpm-solutions by combining process and user-interface modelling. In IFIP Working Conference on The Practice of Enterprise Modeling (pp. 86-97). Springer, Berlin, Heidelberg, 2008
- [8] G.C. Van Der Veer, M. Van Welie, Task Based Groupware Design: putting theory into practice. New York, United States, 2000.
- [9] ISO/IEC 25010: Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models, ISO/IEC JTC1/SC7/WG6, 2011.

GESTIÓN DEL CONOCIMIENTO Y SISTEMAS INFORMÁTICOS. UNA PROPUESTA PARA LAS ORGANIZACIONES DEL SIGLO XXI.

Sonia I. Mariño, María V. Godoy

¹Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura.
9 de Julio 1449. CP: 3400. Corrientes. Argentina. Universidad Nacional del Nordeste,
simarinio@yahoo.com, mvgg2001@yahoo.com

RESUMEN

La compleja sociedad del conocimiento brinda un espacio desde el cual se pueden diseñar, implementar y auditar procesos de gestión de la información para propiciar la administración del conocimiento organizacional como apoyo a la toma de decisiones. Se presenta un proyecto de capacitación y desarrollo en temas de Gestión del Conocimiento vinculado con la Informática, definido desde el ámbito académico y con perspectivas de alcance regional.

Palabras clave: Gestión en la Información, Gestión del Conocimiento, modelos, métodos, herramientas TIC, transferencia de conocimientos.

CONTEXTO

La línea de trabajo se desarrolla en el marco del proyecto “TI en los sistemas de información: modelos, métodos y herramientas” acreditado por la Secretaría General de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE).

1. INTRODUCCIÓN

El conocimiento es un elemento clave en las sociedades del siglo XXI. Este concepto se trata desde los inicios de las sociedades y emerge como un elemento clave en las sociedades complejas. Así, surgen términos como sociedades del conocimiento o sociedades basadas en la economía del conocimiento.

OECD [1, 2] establece que "La economía basada en el conocimiento es una expresión

acuñada para describir las tendencias de las economías avanzadas hacia una mayor dependencia del conocimiento, la información y los altos niveles de habilidad, y la creciente necesidad de acceso fácil a todos ellos por parte del sector empresarial y público”.

Las Universidades como organizaciones complejas del siglo XXI deben afrontar innovadoras estrategias en pro de gestionar sus recursos tangibles e intangibles. El capital humano, el capital estructural y el capital relacional generados en contextos universitarios a diversas escalas ejemplifican claramente el capital intangible organizacional.

En [3, 4] se diferencia entre datos, información, conocimiento y saberes. Por su parte en [5, 6, 7] y otros autores distinguen entre los conocimientos tácitos y explícitos. Este concepto se trata desde diversas disciplinas surgiendo tipologías adecuadas para el tratamiento del conocimiento como clave de la economía basada en el conocimiento.

En este proyecto se propone abordar modelos, métodos y herramientas como un instrumento de la Gestión del Conocimiento y que plasma el Capital Intelectual generado desde un espacio de Educación Superior hacia el contexto.

El objetivo primordial de la Gestión del Conocimiento es propiciar un entorno para crear, transferir y aplicar el conocimiento en las organizaciones [8, 9]. Es decir, se aplica para compartir y difundir el conocimiento y la experiencia de los recursos humanos de una organización con miras a lograr que ésta

sea accesible y empleada por otros miembros de la comunidad.

Los sistemas informáticos reflejan el Capital Humano, el Capital Estructural y el Capital Relacional, que componen el Capital Intelectual de una organización. Dado que surgen de la integración de modelos, métodos y herramientas que pueden tratarse desde la Informática integrando aspectos de la Gestión del Conocimiento.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el marco de este proyecto, las líneas de trabajo relacionadas con los sistemas informáticos y la gestión del conocimiento implican:

- Selección y estudio de Modelos de Gestión del Conocimiento.
- Identificación de métodos y herramientas TIC y no TIC como soporte a los procesos de GC.
- Determinación y estudio de Taxonomías de herramientas TIC y su validación en diversos procesos de desarrollos tecnológicos.
- Diseño de programas de formación de recursos humanos y su seguimiento, en el entendimiento que estos sujetos son los difusores de los conocimientos adquiridos asegurando el alcance regional de las propuestas teóricas-prácticas definidas.

3. RESULTADOS OBTENIDOS / ESPERADOS

El conocimiento adquirido desde un curso de posgrado en el marco de estudios superiores de Ciencias Cognitivas permite identificar las amplias posibilidades de su tratamiento en la Informática, una “disciplina bio-psico-socio-tecno-cultural” según [10].

Los artefactos de las TIC reflejan los modelos, métodos y herramientas seleccionadas por los equipos de trabajo. Es

así que desde la GC representan el capital intelectual de la organización en la cual estos sistemas de apoyo a la administración, gestión y toma de decisiones se insertan. Por ello se plantea establecer una relación entre los procesos de la GC y algunas áreas de la Informática, produciendo innovaciones en la forma en que se trata la captura, almacenamiento, procesamiento y difusión de la información transformada en conocimiento.

Se desarrollan conceptos de GC en trabajos de investigación aplicada y en desarrollos profesionales/ A modo de ilustrar se mencionan los descritos en [11, 12, 13].

Los conocimientos teóricos y desarrollos empíricos logrados se plasman en la definición y ejecución de un curso de posgrado desarrollado en el marco de la Maestría en Tecnologías de la Información (carrera cooperativa entre la Universidad Nacional del Nordeste y la Universidad Nacional de Misiones).

Los resultados alcanzados permiten concretar la socialización de los conocimientos con miras a lograr aprendizajes significativos de modo que la interiorización de estos temas logre mejoras en la gestión de la información y en la definición de sistemas de información destinados a las organizaciones del siglo XXI.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto que se describe se desarrolla implícita y explícitamente desde algunos años en la UNNE.

En referencia a la formación de recursos humanos en la temática se menciona el diseño y ejecución de trabajos de finalización de la carrera Licenciatura en Sistemas de Información, becarios de grado y posgrado, tesistas de posgrado y trabajos de docentes-investigadores.

Además, los conocimientos adquiridos se transfirieron a aproximadamente 80 estudiantes profesionales de la Maestría en Tecnologías de la Información en el ciclo 2016-2017, impartiendo un curso con estos temas.

Los resultados mencionados como la formación de los recursos humanos en torno a temas de la Gestión del Conocimiento acrecentarán la masa crítica de profesionales del sector de las TIC en estos temas que se iniciaron en las Ciencias de la Administración.

5. BIBLIOGRAFÍA

- [1] OECD. The Knowledge-based Economy. OECD. Paris. 1996.
- [2] OECD, "The Measurement of Scientific and Technological Activities: Guidelines for Collecting and Interpreting Innovation Data: Oslo Manual, Third Edition" prepared by the Working Party of National Experts on Scientific and Technology Indicators, OECD, Paris, 2005.
- [3] K. M. Wiig, K. *Knowledge management foundations*. Arlington, TX: Schema Press, 1993.
- [4] K. M. Wiig, *Integrating Intellectual Capital and Knowledge Management. Long Range Planning*, 30(3) pp. 372, 1997.
- [5] M. Polanyi, *The tacit dimension*, 1st edition, Garden City, New York: Doubleday and Company, 1966.
- [6] I. Nonaka y H. Takeuchi, "The knowledge-creating company: How Japanese companies create the dynamics of innovation", *Harvard Business Review*, 69 (6), pp. 96-104, 1999.
- [7] T. Davenport, L. Prusak, *Working Knowledge*. Harvard Business School Press, Boston, Massachusetts, 1998.
- [8] M. Alavi, D. Leidner, *Sistemas de gestión del conocimiento*. España: Thomson, 2002.
- [9] A. Pérez Lindo, C. Varela, F. Grosso, C. Camós, A.M. Trottni, M.L. Burke, y S. Darin, *Gestión del conocimiento: un nuevo enfoque aplicable a las organizaciones y la universidad*, Ed. Grupo Editorial Norma, 2005.
- [10] G. E. Barchini, "Informática. Una disciplina bio-psico-socio-tecnocultural", *Revista Ingeniería Informática*, 12, pp. 1-12, 2006.
- [11] M. V. Godoy, S. I. Mariño, P. L. Alfonzo, M. Cáceres, R. Alderete, "Sub-redes de gestión y sus herramientas como apoyo en un Departamento Pedagógico Universitario", *XIX Workshop de Investigadores en Ciencias de la Computación*, 2017.
- [12] S. I. Mariño, "Los sistemas expertos para apoyar la gestión inteligente del conocimiento", *Vinculos*, 11(1), pp. 101-108, 2014.
- [13] S. I. Mariño, P. L. Alfonzo, M. V. Godoy, Sistematización de trabajos de graduación, una aproximación a la gestión del conocimiento, *EDUTECH, Revista Electrónica de Tecnología Educativa*, 57, pp. 33-44, 2016.

FASES PARA EL DESARROLLO DE PROGRAMAS PARALELOS Y SU APLICACIÓN EN LA PREDICCIÓN DEL COMPORTAMIENTO DE INCENDIOS FORESTALES

Ana Laura Molina¹, Nelson Rodriguez¹, María A. Murazzo¹, Fernando Pincioli² y Miguel Méndez-Garabetti³

¹Departamento de Informática, Facultad de Ciencias Exactas, Físicas y Naturales
Universidad Nacional de San Juan

²Instituto de Investigaciones, Facultad de Informática y Diseño
Universidad Champagnat

³Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

lauramolina@outlook.com, nelson@iinfo.unsj.edu.ar, marite@unsj-cuim.edu.ar,
pincirolifernando@uch.edu.ar, mendez-garabettimiguel@uch.edu.ar

RESUMEN

En la actualidad, la programación paralela está siendo ampliamente utilizada en aplicaciones tanto científicas como comerciales. A través de esta se busca lograr una reducción del tiempo necesario para resolver un problema complejo y un aumento en la calidad de la respuesta. Las propuestas existentes para guiar el desarrollo de programas paralelos plantean únicamente metodologías para su diseño y en algunos casos herramientas para su implementación, sin hacer uso de buenas prácticas propias de la Ingeniería de Software.

Esta línea de investigación intenta elaborar una propuesta de fases que incluirá por una parte las de diseño e implementación -ya investigadas previamente por diversos autores-, y por otra parte la adaptación de fases propias de la Ingeniería de Software como la captura de requisitos y las pruebas, no utilizadas hasta el momento en la programación paralela. Esta propuesta será aplicada a la predicción del comportamiento de incendios forestales, donde el poder computacional ofrecido por los programas paralelos contribuiría en la obtención de un resultado de calidad en el menor tiempo posible.

Palabras clave: programas paralelos, Ingeniería de Software, predicción de incendios forestales.

CONTEXTO

El presente trabajo se encuadra en la línea de I+D denominada “Procesamiento Distribuido y Paralelo” y se desarrolla en el marco del proyecto de investigación denominado: “Orquestación de servicios para la Continuidad Edge al Cloud”.

La unidad ejecutora de dicho proyecto es el Departamento de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales (FCEfyN) de la Universidad Nacional de San Juan (UNSJ).

1. INTRODUCCIÓN

En un mundo donde el tiempo es, quizás, el bien más preciado, la velocidad resulta esencial y la eficiencia es traducida en cuán rápido y qué tan bien podemos solucionar un problema [1]. En este contexto, es donde el procesamiento paralelo adquiere importancia. Esta, es un área de las ciencias de la computación que aprovecha los recursos de hardware a través de la construcción de programas en los que múltiples tareas

cooperan estrechamente para resolver un problema [2].

Un programa paralelo divide el problema a resolver en piezas que serán ejecutadas por un procesador multi-núcleo o múltiples procesadores mono-núcleo, dependiendo de si se trata del modelo de memoria compartida o distribuida. Algunas de estas piezas lo harán en simultáneo, es decir en paralelo, logrando así una reducción del tiempo necesario para completar el trabajo [3]. Además de mejorar el tiempo de ejecución, también se puede aumentar la calidad de la solución obtenida respecto del mejor algoritmo secuencial tratando el mismo problema [1]. Sin embargo, no todas las aplicaciones que sean programadas en paralelo obtendrán una importante mejora en el desempeño. Esto se debe, en la mayoría de los casos, a la necesidad de los núcleos de coordinar su trabajo. Esta coordinación involucra comunicación, balanceo de carga y sincronización [2].

La evolución de los procesadores a lo largo del tiempo, demostró que existe una continua demanda de incremento del poder computacional. Otra fuente histórica de incremento del poder computacional es el paralelismo que con el paso de los años, se está volviendo omnipresente y, en consecuencia, la programación paralela se convierte así en un aspecto central para las empresas de desarrollo de software.

A medida que aumenta el poder computacional, aumenta la cantidad de problemas que pueden ser considerados para su resolución. Entre estos se encuentran los pertenecientes a campos como: modelado climático, descubrimientos medicinales, análisis de datos en genética, física, medicina, Internet, entre otros [2]. Dentro del conjunto de problemas que pueden ser resueltos a partir de un incremento del poder de cómputo, existe un subgrupo que requiere su resolución en un período de tiempo razonable y otro que incluso posee un tiempo máximo de resolución, denominado en inglés *deadline*, donde posterior a este, el dato deja de ser útil, como es el caso de la predicción meteorológica [4] y la predicción del

comportamiento de incendios forestales [5][6][7]. En este último caso, para poder tomar decisiones preventivas sobre una distribución de recursos de extinción lo más eficiente posible, se necesita un sistema de predicción cuyos resultados deben producirse antes de que ocurra el fenómeno [5]. Por esto, se vuelve fundamental lograr una reducción en el tiempo de cálculo. Si bien el tiempo de respuesta es clave, también lo es la calidad de la predicción debido a la criticidad de las decisiones que se van a tomar en función de la salida [6]. En este caso se busca determinar el posible comportamiento de un incendio forestal una vez que este ya se ha iniciado, es decir, obtener información de la dirección y velocidad de la propagación, la forma del frente de fuego y la intensidad de la llama, utilizando la mayor cantidad de información posible [7][8]. De acuerdo con estos requisitos, se torna primordial la utilización de sistemas de procesamiento de alto rendimiento y la implementación de programas paralelos [7]. Esta línea de investigación se basa en la aplicación de los programas paralelos a la predicción del comportamiento de incendios forestales.

Si bien, tradicionalmente, los principales desarrollos de programas paralelos estuvieron orientados a aplicaciones científicas e ingenieriles, actualmente, las aplicaciones comerciales que deben procesar grandes volúmenes de datos también requieren del paralelismo para obtener un buen desempeño.

En cualquiera de los casos mencionados, desarrollar programas paralelos no es una tarea sencilla, ya que la mayoría de los problemas tienen muchas soluciones paralelas y resulta difícil encontrar la óptima o, incluso, una aceptable. De hecho, la programación paralela ha sido pensada tradicionalmente como una actividad intensiva en tiempo y esfuerzo [9]. Tal es así que el progreso de los componentes de hardware implicados en la computación paralela, debe ir acompañado de un progreso en los aspectos de diseño, análisis y aplicación de los programas paralelos para que estos beneficios sean observables en la práctica [1].

Tradicionalmente, la Ingeniería de Software ha estado más enfocada en el dominio empresarial y de la tecnología de la información que en el dominio de la ciencia y la ingeniería. Muchos investigadores pertenecientes a esta rama de la Ingeniería han asumido que cualquiera de sus soluciones podría aplicarse al ámbito de las ciencias de la computación, mientras que un gran número de investigadores de dicho ámbito han creído que sus inconvenientes eran diferentes a los surgidos en entornos tradicionales de la Ingeniería de Software y, por lo tanto, no han aplicado sus métodos.

Actualmente, en la mayoría de los casos, el desarrollo de software en las ciencias de la computación no sigue las buenas prácticas tradicionales de la Ingeniería de Software y esto se debe a que muchas de ellas no son aplicables sin una previa adaptación. Esta tendencia debe ser abordada para mejorar tanto la productividad del desarrollo como la fiabilidad del software a desarrollar [10]. No tendría sentido buscar notaciones, métodos o técnicas universales para la Ingeniería de Software ya que diferentes tipos de software requieren de distintos enfoques. Sin embargo, hay ciertos principios fundamentales que se aplican a todo tipo de software como son: desarrollar utilizando un proceso comprendido y gestionado, entender y gestionar los requerimientos de clientes y usuarios, y usar eficientemente los recursos [11].

A partir del análisis de propuestas de diversos autores [1][2][9][12], es posible afirmar que todos ellos plantean únicamente metodologías para el diseño de programas paralelos y en algunos casos herramientas para la implementación, dejando de lado las fases frecuentemente utilizadas en el área de la Ingeniería de Software, como son la captura de requisitos y las pruebas.

Estas dos últimas fases mencionadas se presentan como un espacio sin investigar en el ámbito de la programación paralela, por lo cual se transforma en el problema que se intenta abordar en esta línea de investigación. El problema planteado se relaciona específicamente con la conveniencia de

integrar aportes de la Ingeniería de Software tradicional al desarrollo de programas paralelos, reuniendo conocimientos hasta el momento inconexos. Esta investigación se propone definir y caracterizar las fases para el desarrollo de programas paralelos y aplicar la propuesta a la predicción del comportamiento de incendios forestales.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Siguiendo la línea de investigación en que se enmarca este trabajo, se llevarán a cabo actividades relacionadas con los siguientes espacios:

- Ingeniería de software
- Ciencias de la computación, en el área de Procesamiento Distribuido y Paralelo.

3. RESULTADOS OBTENIDOS/ESPERADOS

En esta línea de I+D se intenta elaborar una propuesta de fases que incluya tanto las de diseño e implementación que suelen utilizarse frecuentemente en el ámbito de la programación paralela como la adaptación de fases propias de la Ingeniería de Software, como son la captura de requisitos y las pruebas de verificación y validación, que no se han utilizado aún en dicho ámbito.

Estas dos últimas fases serán adaptadas en función de las necesidades específicas y características particulares de los programas paralelos y sus entornos de desarrollo. Para llevar a cabo esta tarea, se seleccionarán entre las diversas técnicas existentes y difundidas, aquellas que mejor se ajusten, y serán modificadas en aspectos que se considere que contribuirán a mejorar la calidad del software resultante.

Se busca contribuir en el área de la programación paralela produciendo una guía útil para que quienes no tienen conocimientos en el área de la Ingeniería de Software puedan llevar a cabo dichas fases adecuadamente. Para cada fase de la propuesta se plantearán

una serie de actividades a realizar y un conjunto de buenas prácticas que deberían tener en cuenta las partes interesadas tanto pertenecientes al equipo de desarrollo como ajenas (clientes, usuarios, sponsors, etc.).

Luego, se pretende aplicar la propuesta de fases elaborada al desarrollo de un programa paralelo para la predicción del comportamiento de incendios forestales y así poder analizar sus limitaciones y potencialidades.

Se espera que a través del uso de la propuesta en el desarrollo de programas paralelos se logre satisfacer las necesidades y objetivos de todas las partes interesadas.

4. FORMACIÓN DE RECURSOS HUMANOS

Esta investigación se encuadra en el proyecto que se ejecuta bajo la dirección del Lic. Nelson Rodríguez y la codirección de la Lic. María A. Murazzo.

En el marco de la línea de I+D presentada, se desarrolla una tesina de grado por parte de la estudiante de la Licenciatura en Sistemas de Información de la Universidad Nacional de San Juan, Ana Laura Molina.

Dicha tesina es asesorada por el director del proyecto antes mencionado, el Lic. Fernando Pinciroli y el Mgter. Ing. Miguel Méndez-Garabetti.

5. REFERENCIAS

- [1] S. G. Akl y M. Nagy, "Introduction to Parallel Computation" en *Parallel Computing: Numerics, Applications, and Trends*, R. Trobec, M. Vajteršic and P. Zinterhof, London: Springer, 2009, pp. 43-80.
- [2] P. Pacheco, *An introduction to parallel programming*. USA: Elsevier, 2011.
- [3] A. Freeman, *Pro. NET 4 parallel programming in C*. USA: Apress, 2011.
- [4] B. Wilkinson, y M. Allen, *Parallel programming: techniques and applications using networked workstations and parallel computers*, 2nd ed. New York: Prentice hall, 2005.
- [5] M. Méndez-Garabetti, G. Bianchini, M. L. Tardivo, y P. Caymes-Scutari, "Comparative analysis of performance and quality of prediction between ESS and ESS-IM", *Electronic Notes in Theoretical Computer Science*, vol. 314, pp. 45-60, 2015.
- [6] M. Méndez-Garabetti, G. Bianchini, M. L. Tardivo, P. Caymes Scutari y G. V. Gil Costa, "Hybrid-Parallel Uncertainty Reduction Method Applied to Forest Fire Spread Prediction", *Journal of Computer Science & Technology*, vol. 17, pp. 12-19, 2017.
- [7] M. Mendez-Garabetti, M. L. Tardivo, G. Bianchini, P. Caymes Scutari, "Predicción del Comportamiento de Incendios Forestales mediante un Método de Reducción de Incertidumbre basado en HPC y Evolución Diferencial", en *XVI Workshop de Investigadores en Ciencias de la Computación*, 2014, pp. 690-694.
- [8] M. Mendez-Garabetti, G. Bianchini, P. Caymes-Scutari, M. L. Tardivo, "Increase in the quality of the prediction of a computational wildfire behavior method through the improvement of the internal metaheuristic", *Fire Safety Journal*, vol. 82, pp. 49-62, 2016.
- [9] A. Grama, A. Gupta, G. Karypis, y V. Kumar, *Introduction to Parallel computing*. Harlow, UK: Pearson, 2003.
- [10] J. C. Carver, "Software Engineering for Computational Science and Engineering", *Computing in Science & Eng*, vol. 14, no. 2, pp. 8-11, 2012.
- [11] I. Sommerville, *Software Engineering*. Boston: Addison Wesley Publishing Company, 2011.
- [12] I. Foster, *Designing and building parallel programs*. Boston: Addison Wesley Publishing Company, 1995.

MODELO DE SISTEMA DE GESTION DE CALIDAD PARA UN INSTITUTO DE EDUCACION A DISTANCIA

Jorge A. Silvera, Daniel Arias Figueroa, Gustavo Gil, Valeria González

Centro de Investigación y Desarrollo en Informática Aplicada (C.I.D.I.A.)

Facultad de Ciencias Exactas – Universidad Nacional de Salta – Salta - Argentina

jsilvera@unsa.edu.ar, gdgil@unsa.edu.ar, daaf@cidia.unsa.edu.ar,

ygonzalez@cidia.unsa.edu.ar

RESUMEN

El presente trabajo propone un modelo de software resultado de la aplicación de la metodología de desarrollo de software Webml y el estándar IFML aplicado a la gestión de la calidad del Instituto de Investigación y Educación a Distancia (I.I.E.Di.), dependiente de la Secretaria Académica de Rectorado de la Universidad Nacional de Salta (U.N.S.a).

Palabras claves:

Normas de Calidad ISO 9.001, Sistema de Gestión de Calidad (SGC), Institutos de Investigación, Herramientas informáticas de apoyo al SGC. Metodología WebML (Web Modeling Language). IFML (Interaction Flow Modeling Language).

CONTEXTO

El Instituto de Investigación y Educación a Distancia (I.I.E.Di.) fue creado en el año 2.002 por el Consejo Superior mediante Resolución CS N° 037/02, dependiendo en ese momento de la Facultad de Ciencias de la Salud de la Universidad Nacional de Salta. En el año 2.006 pasa a depender de la Secretaria Académica de Rectorado – U.N.S.a por resolución CS N° 349/06.

Entre los objetivos generales del I.I.E.Di. podemos mencionar:

- a) Promover y realizar investigaciones en educación a distancia.
- b) Propender a la formación y capacitación de recursos humanos en distintas disciplinas en la modalidad a distancia y

en el uso de Tecnologías de la Información y la Comunicación para coadyuvar a la calidad de las propuestas a distancia.

- c) Asesorar y/o prestar servicio en el ámbito de la modalidad a distancia a las distintas Unidades Académicas y Sedes Regionales de la U.N.Sa. y entidades públicas y/o privadas.
- d) Difundir las acciones desarrolladas por este Instituto.
- e) Promover el uso de medios y nuevas tecnologías en la enseñanza presencial.
- f) Producir materiales en distintos soportes para educación a distancia y enseñanza presencial.
- g) Impulsar la Educación a distancia en entornos virtuales en la U.N.Sa y fijar pautas, criterios y procedimientos que garanticen su calidad

Dentro de las actividades que desarrolla podemos destacar el Proyecto de Tutorías Personalizadas, cuyo propósito es realizar acompañamiento a equipos docentes interesados en la inclusión de aulas virtuales en las cátedras universitarias.

El presente trabajo se lleva a cabo dentro del marco del proyecto de investigación Nro: 2.278/0 “Estudio de la familia de normas ISO 9.000 y su aplicación a centros educativos”, aprobado en el año 2.014 por el consejo de investigación de la Universidad Nacional de Salta.

1. INTRODUCCION

El presente trabajo fue llevado a cabo considerando dos puntos de vistas bien diferentes, por un lado, se estudian y definen directrices o guías que orienten a los institutos de investigación y capacitación a distancia, en cuanto a la implementación de un sistema de gestión de calidad (SGC) eficaz que cumpla los requisitos de la norma ISO 9001:2015. Por otro lado, se plantea el análisis, diseño e implementación de un software de apoyo al SGC para el Instituto de Investigación y Educación a Distancia (I.I.E.Di.), utilizando como metodología Web Modeling Language (WebML) y la herramienta CASE WebRatio, expresando los modelos en el standard IFML (Interaction Flow Modeling Language). Se espera que la presente investigación sirva de referencia para cualquier implementación de gestión de la calidad en el ámbito de la capacitación a distancia.

A modo de breve reseña histórica, en el año 1.987 se obtuvo la primera referencia europea a la calidad en investigación y desarrollo (I+D), que analiza la aplicabilidad de la norma ISO 9.001:2.008 a las actividades de I+D [Alonso, P. M. 2.005].

La implementación exitosa de un sistema de calidad (SGC) aporta un gran número de beneficios a las organizaciones en general, logrando no solo reducir sus costos de manera razonable, sino que además ayudan a lograr la tan preciada satisfacción de sus usuarios, lo cual nos brinda una gran motivación para los integrantes de la misma.

Promover la calidad en la capacitación e investigación a distancia es tratar de mejorar de forma continua sus prácticas de forma que permitan:

- Garantizar los resultados y productos obtenidos.
- Asegurar la trazabilidad de los procesos que se ejecutan en la organización.

Desde el momento de su creación, los requisitos de la norma internacional ISO 9.001:2015 son genéricos y aplicables a todas las organizaciones sin importar su tipo, tamaño o producto o servicio suministrado.

La gestión de la calidad en la investigación debe ser sobre todo flexible y adaptada a las necesidades específicas de los investigadores, pudiendo conceptualizarse tres fases asociadas a la investigación:

- 1) Definición de los objetivos: En esta fase hay que lograr la identificación de los criterios de satisfacción de las diferentes partes interesadas.
- 2) Realización de la investigación: Las cuestiones a tener en cuenta en esta fase tienen que ver con los diferentes procesos a ejecutar, sus interacciones y gestión eficiente.
- 3) Valoración y puesta en valor de los resultados: Esta fase es esencial pues contribuye al reconocimiento del organismo de investigación y en ella se concretan los esfuerzos y la razón de ser de los investigadores.

La representación de las tres fases principales de la investigación junto con los principios de gestión de calidad aplicables se muestra en la siguiente figura:

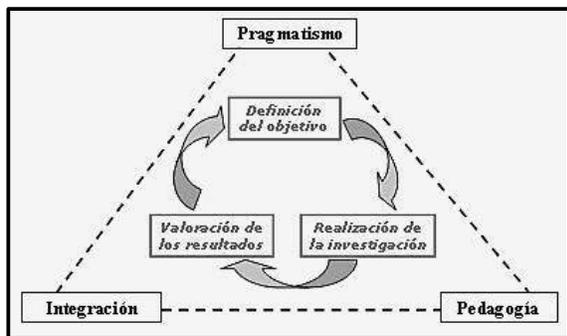


Figura 1- Fases principales de la investigación.

Podemos definir una metodología para implementar el SGC en el área de la capacitación e investigación a distancia, utilizando el esquema anterior de manera sistemática, dando lugar a un ciclo de mejora continua, tal y como se representa en la siguiente figura:

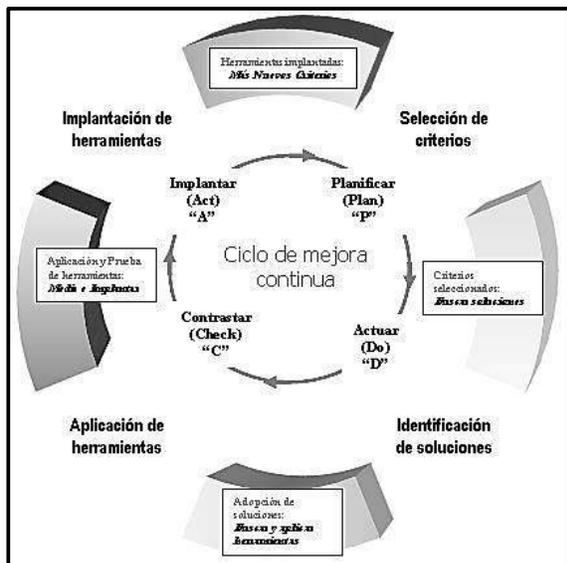


Figura 2- Ciclo de mejora continua.

Como hemos podido observar, en base a la investigación bibliográfica, un sistema de gestión de la calidad se basa en la gestión de muchos documentos, de manera estricta y minuciosa. Por esta razón es fundamental el apoyo de la tecnología informática, es decir que no alcanza con un procesador de textos y un espacio de almacenamiento compartido en un servidor de archivos.

Todo documento relacionado con el sistema de gestión de la calidad debe ser desarrollado a través de un proceso perfectamente documentado. Además, una vez aprobado, debe estar disponible para todos quienes participan en el SGC, y por otra parte no disponible para quienes no están involucrados en el. A su vez las sucesivas revisiones deben quedar claramente identificadas así como los cambios realizados. También, si los documentos obsoletos se mantienen en el SGC para poder ser consultados deben quedar claramente identificados como obsoletos para impedir que sean utilizados como actuales, por error. Por ejemplo, debe impedirse su modificación incluso a personal autorizado para generar documentos.

Esto es claramente el manejo de una base documental asociada a un proceso de decisión y de elaboración conjunta de los que típicamente se realizan con herramientas de trabajo en grupos.

En función a lo expresado, tomando como referencia las especificaciones y directrices investigadas por el C.I.D.I.A. para la aplicación de la familia de norma ISO 9.001:2.008 se realizó el análisis y diseño de un SGC, utilizando la metodología WebML, logrando expresar los modelos en el standard IFML (Interaction Flow Modeling Language) y se implementó de un prototipo funcional de una herramienta para acompañar una implementación de calidad bajo los requisitos de la norma ISO 9.001:2.015 en el Instituto de Investigación y Educación a Distancia (I.I.E.Di.). Posteriormente el prototipo se convirtió en producto final.

Para la especificación de requerimientos de calidad se utilizó el estándar ISO/IEC 25030:2007 que es parte del standard internacional SQuaRE, de esta manera los requisitos de usuarios son expresados como requisitos funcionales.

El enfoque de la metodología WebML combina componentes tradicionales bien conocidos por los desarrolladores, como el

diseño conceptual de datos usando el modelo Entidad Relación y la especificación de los casos de usos usando UML, con nuevos conceptos y métodos para el diseño de hipertextos, que son fundamentales para el desarrollo web. No obstante, el valor del enfoque de la propuesta no está en los componentes individuales, sino en la definición de un marco sistemático para que las actividades de desarrollo de aplicaciones web puedan ser organizado de acuerdo a los principios fundamentales de la ingeniería de software, de modo que todas las tareas encuentren un soporte adecuado a partir de los conceptos, notaciones y técnicas propuestas por las metodologías. La característica distintiva de este marco de desarrollo es el énfasis en el modelo conceptual, el cual ha sido probado con éxito en muchos campos de la ingeniería del software, y en diseño de base de datos, donde el modelo Entidad Relación ofrece una notación de alto nivel e intuitiva para la comunicación de los requisitos de información entre los diseñadores y no técnicos, y es la base para la creación de esquemas de bases de datos de alta calidad.

En esencia, WebML consiste en gráficos simples y conceptuales para expresar un hipertexto como un conjunto de páginas y operaciones, WebML representa una página como una estructura compuesta por unidades de contenido y links.

WebML propone que un sitio web conste de tres grandes partes conceptuales, la estructura, el hipertexto y la presentación.

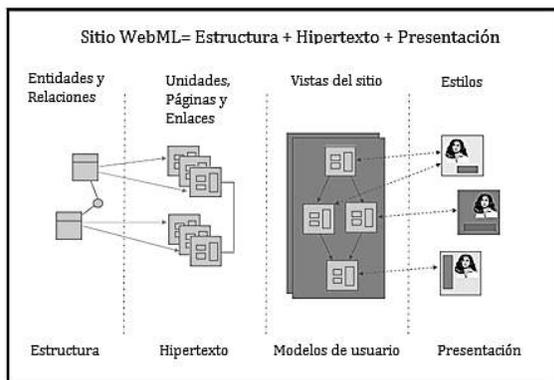


Figura 3- Conceptos principales de WebML.

La metodología está compuesta por la creación de los Modelos de Datos, Hipertexto, Presentación y Personalización, expresando los modelos en el standard IFML (Interaction Flow Modeling Language). En este trabajo se desarrollaron todos los modelos de la metodología y para la implementación se utilizó la herramienta WebRatio con licencia de uso gratuito para el modelado de procesos de negocios (BPM).

En síntesis, los resultados obtenidos a la fecha son los siguientes:

- El Análisis y diseño de un Sistema de Gestión de Calidad para el I.I.E.Di con modelos expresados en el standard IFML (Interaction Flow Modeling Language), *junto* a la especificación de requerimientos, aplicados a los procesos específicos relacionados a la capacitación e investigación a distancia.
- Un prototipo funcional de un Sistema de Gestión de Calidad para institutos de educación e investigación a distancia.

3. LINEAS DE INVESTIGACION y DESARROLLO

Los principales ejes temáticos que se están investigando son los siguientes:

- ✓ Tecnología Informática aplicada en Educación.
- ✓ Gestión de Calidad aplicada a Institutos de Investigación Universitaria.
- ✓ Herramientas informáticas para la implementación de un SGC ISO 9.001.
- ✓ Aplicación de la metodología WebML con IFML para el diseño de una herramienta que apoye a la implementación de SGC ISO 9.001.

4. FORMACION DE RECURSOS HUMANOS

La estructura del equipo de investigación es de 4 (cuatro) miembros incluidos el Director y Co-director.

Uno de sus miembros obtuvo la Especialidad en Ingeniería de Software de la Universidad Nacional de La Plata, con el trabajo “Sistema de Gestión de Calidad bajo Normas ISO”.

Otro de sus miembros alcanzó el título de Licenciado en Análisis de Sistemas, otorgado por la Universidad Nacional de Salta, con la tesis “Análisis y Diseño de un Sistema de Gestión de Calidad basado en la Norma ISO 9.001”.

Además, otro de sus miembros continúa realizando el trabajo final del Magíster en Administración de Negocios de la Universidad Católica de Salta en el área de las normas de calidad.

Continuamos con la dirección de tesis de grado de alumnos de la carrera Licenciatura en Análisis de Sistemas de la Universidad Nacional de Salta.

4. RESULTADOS OBTENIDOS/ESPERADOS

En el presente trabajo se ha abordado la gestión de los sistemas de calidad desde una perspectiva de los centros de capacitación e investigación a distancia.

Consideramos que, además de todas las herramientas disponibles, la aplicación de la familia de normas ISO 9001:2015 junto con el apoyo de un modelo adecuado de software, con la especificación de requerimientos adecuado, se constituye en una estrategia importante para alcanzar la satisfacción de los alumnos a distancia.

Actualmente se está considerando la posibilidad de alcanzar la certificación del Instituto de Investigación y Educación a Distancia (I.I.E.Di.), por algunas de las entidades certificantes, tales como Bureau Veritas Quality International e IRAM, las cuales son las más conocidas en nuestro país.

5. BIBLIOGRAFIA

- Arias Figueroa, Daniel y otros (2.011). Normas ISO y su Aplicación en Centros Educativos. XIII Workshop de Investigadores en Ciencias de la Computación. ISBN 978-950-673-892-1.
- Ishikawa, Kaoru (1.997). ¿Qué es el control de la Calidad?: La modalidad Japonesa. Bogotá: Grupo Editorial Norma.
- Fernández, A. (2.006.). Implantación de un sistema de gestión de la calidad Norma ISO 9.001:2.001. Centro de la Calidad de Asturias/ Instituto de Fomento Regional.
- Norma ISO 9.001 (2.015) Elaborada por el Comité Técnico ISO/TC176 de ISO.
- Guía de Interpretación de la IRAM-ISO 9.001:2008 para la educación. IRAM 30.000:2.001.
- Garbarini, R., Cigliuti, P., Burstyn, A., Pollo-Cattaneo. (2.013). Implementación de un Sistema de Gestión de Calidad y Servicios en Laboratorio Universitario de Ingeniería en Sistemas de Información Implementación de un Sistema de Gestión de Calidad y Servicios en Laboratorio Universitario de Ingeniería en Sistemas de Información. VIII Congreso de Tecnología en Educación y Educación en Tecnología. Quilmes. ISBN 978-987-1676-04-0.
- Orthustegui, F. y otros. (2.011). Análisis del Impacto de la Implementación de un Sistema de Gestión de la Calidad Interno para los Laboratorios de Enseñanza e Investigación de la UNLaM. XXXVI Jornadas IRAM Universidades – XXIII - Foro UNILAB. UNSL.
- Alonso, P. M. (2.005). Revista Madri+d, Número 32. Calidad En Investigación 1ª Parte. De qué trata la Gestión de Calidad en Investigación.

- Alonso, P. M. (2.005). Calidad En Investigación 2ª Parte. Aproximación metodológica a la mejora de las actividades de investigación. Revista Madri+d, Número 33.
- Web Modeling Language, sitio oficial de la Metodología <http://www.webml.org>.
- WebRatio, sitio oficial de la herramienta <http://www.webratio.com>.
- IFML <http://www.ifml.org>
- Centro de Investigación y Desarrollo en Informática Aplicada, sitio oficial <http://cidia.unsa.edu.ar>.
- Norma ISO/IEC 25010 (2.007). Software Engineering–Software Product Quality Requirements and Evaluation (SQuaRE) – Quality model and guide elaborada por el Comité Técnico ISO/IEC JTC 1/SC 7 de ISO.

Modelo para la interoperabilidad entre controladores de Redes Definidas por Software.

Juan Carlos Calloni
Universidad Tecnológica
Nacional F.R. San
Francisco
jcalloni@hotmail.com

**Daniel Fernandes
Macedo**
Universidad Federal
de Minas Gerais
damacedo@dcc.ufmg.br

**Germán A.
Montejano**
Universidad Nacional
de San Luis
germanamontejano@gmail.com

**Vinicius Fonseca e
Silva**
Universidad Federal
de Minas Gerais
viniciusfs@dcc.ufmg.br

Erik de Britto e Silva
Universidad Federal
de Minas Gerais
erikbritto@gmail.com

Eduardo Scarello
Universidad Tecnológica
Nacional F.R. San
Francisco
eduardo.scarello@gmail.com

Federico Degiovanni
Universidad Tecnológica
Nacional F.R. San
Francisco
federicoadegiovanni@gmail.com

Paez Sergio
Universidad Tecnológica
Nacional F.R. San
Francisco
sergiopaez1000@gmail.com

Resumen

Software Defined Networking (SDN) o las redes definidas por software, se enfocan en la programación por software de las mismas, en el cual el control se desvincula del hardware. El plano de control es separado de la capa de red física y puede controlar flujos por separado dependiendo de las políticas definidas.

El Controlador es una parte importante de la red SDN, que actúa como un cerebro virtual. No sólo puede monitorizar el tráfico de una red con facilidad, sino que le ordena a los sistemas por debajo, como switches, routers y otros equipos de la red, de cómo manejar el tráfico, haciendo una gestión inteligente, cumpliendo con las políticas que se programaron en el mismo.

Pero si dos dominios SDN se quisieran comunicar entre sí para distribuir sus políticas, como priorizar paquetes en el caso de video conferencia para mejorar la calidad de servicio, o en el caso de paquetes multimediales a los cuales se les quiera dar mayor prioridad. Los controladores SDN podrían establecer relaciones de confianza unidireccional o bidireccional entre esos paquetes de forma dinámica. Para ello se debería contar con un modelo única para que los diferentes controladores para que puedan comunicarse o Interoperar entre sí.

Palabras Clave

Redes definidas por Software, Modelo, OpenFlow, Interoperabilidad, calidad de servicio, Controladores.

Contexto

La línea de investigación presentada se encuentra en el marco de las áreas prioritarias para el desarrollo de las actividades de I+D de Ingeniería de Software entre la Universidad Nacional de San Luis, Facultad de Cs. Físico Matemáticas y Naturales de la Ingeniería en Informática en conjunto con la Universidad Federal de Minas Gerais y su laboratorio de redes Winnet y la Universidad Tecnológica Nacional Facultad Regional San Francisco con la carrera de Ingeniería en Sistemas de Información y el Grupo GarLan. En ese marco se vienen llevando actividades en conjunto a través del Programa Binacional de Centros Asociados para el Fortalecimiento de Posgrados Brasil-Argentina (CAFP-BA) aprobada por la Secretaria de Políticas Universitarias (SPU) de Argentina y la Fundação Coordenação de Aperfeiçoamento do Pessoal de Ensino Superior (CAPES) de Brasil.

El área principal es Ingeniería de Software relacionándose con Redes definidas por Software. Éstas líneas se encuentran insertas en el Proyecto de Investigación y Desarrollo (PID) Nro SIUTNSF0004906 “Redes Definidas por Software modelo para la interoperabilidad de controladores de diferentes dominios”, el mismo se encuentra

en desarrollo. Dicho proyecto es ejecutado por el grupo de I+D GARLAN de la UTN Facultad Regional San Francisco.

El proyecto se encuentra homologado y financiado por la Secretaría de Ciencia y Tecnología de la Universidad Tecnológica Nacional, y está incluido en el Programa I&D + i Sistemas de Información e Informática de la Universidad Tecnológica Nacional creada por Res. CSU. 2508/16 y Disp. SCyT N° 336/2016.

1. Introducción

Las redes definidas por software SDN, básicamente se enfocan en la programación por software de las redes en donde el control se desvincula del hardware. El plano de control es separado de la capa de red física y puede controlar flujos por separado, dependiendo de las necesidades y de las políticas en capas superiores [1].

Un controlador SDN actúa como un cerebro virtual de la red, y ofrece a los administradores una vista general de la red. No sólo puede monitorizar el tráfico de una red con facilidad, sino que ordena a los sistemas por debajo, como switches, routers y otros equipos de la red, cómo deben manejar el tráfico de red, haciendo una gestión inteligente del mismo. En definitiva el controlador escribe directivas en las tablas de flujo de los conmutadores openflow, haciendo que los dispositivos individuales relacionados a las redes tradicionales desaparezcan bajo el nuevo paradigma de las SDN. En otras palabras un conmutador openflow puede ser de acuerdo a las reglas escritas en las tablas de flujo, un switch, un router, o un firewall o lo que las reglas definan. Este enfoque de las redes puede responder fácilmente a las cambiantes necesidades del negocio y dar forma al flujo del tráfico, sin tener que buscar ni manipular equipos o hardware individuales [2].

Para que esto sea posible se utiliza el protocolo OpenFlow que toma las decisiones de envío de paquetes de forma centralizada, haciendo posible la programación de la red. Nace como protocolo experimental después de 6 años de investigación entre la Universidad de Stanford y Berkeley. Es un protocolo que permite a un

servidor Controlador SDN, comunicarle a la red hacia dónde mandar los paquetes y cómo tratarlos a través de las tablas de flujos que existen en los Switch OpenFlow. OpenFlow fue diseñado para la programación de redes y tiene como objetivo permitir que funcionen las SDN. A principios de 2012, la red interna de Google funcionaba completamente en OpenFlow [3]. En el año 2014 Google terminó de implementar B4, una WAN privada que conecta los centros de datos de Google en todo el planeta a través de SDN [4].

En junio de 2014, Facebook anunció, el nuevo switch "Wedge" de top-of-rack (TOR), junto con un sistema operativo Linux llamado "FBOSS". A diferencia de los tradicionales Switch de hardware cerrado, con 'Wedge' cualquiera puede modificar o reemplazar cualquiera de los componentes de nuestro diseño para satisfacer mejor sus necesidades, dijo Facebook [5].

Con el fin de escalar la red para satisfacer las necesidades de las aplicaciones y aumentar el tráfico de máquina a máquina, la estructura tradicional de la red basada en la jerarquía no es muy buena y por eso Facebook decidió implementar su propio SDN. [6]

Otro escenario donde SDN es muy importante es en RNP (Rede Nacional de Pesquisa de Brasil) La Red Nacional de Educación e Investigación (RNP). [7]

Cuando hablamos de SDN y mencionamos que en la capa más importante de estas redes hay software, entonces aquí es donde la Ingeniería de software pasa a tener un rol fundamental en el diseño, gestión y administración de las redes. Ahora las redes no van a depender por debajo de la capa física y del fabricante y sus diferentes firmwares o sistemas operativos, sino que dependerán de un controlador programado con un software. En donde aparece un protocolo común a todos los dispositivos para la comunicación entre el controlador SDN y la capa de infraestructura, denominado OpenFlow [8].

Los dominios SDN con diferentes controladores, a veces necesitarían comunicarse entre sí, por ejemplo, para garantizar la calidad del servicio. Ahora bien, si dos dominios SDN se quisieran comunicar

entre sí para distribuir sus políticas, como puede ser, la de priorizar paquetes en el caso de videos conferencias, en el caso de paquetes multimediales, para un auto autónomo, para telemedicina o para streaming multimedia bajo demanda, en todos estos casos lo que se busca es mejorar la calidad de servicio y para ello no hemos encontrado nada que haga que comuniquen sus políticas dos controladores SDN.

En esos casos, los diferentes controladores SDN de distintos dominios, podrían establecer relaciones de confianza unidireccional o bidireccional entre esos paquetes de forma dinámica. Entonces los software SDN deberían contar con un modelo o una ontología única para que los controladores puedan comunicarse o interoperar entre sí. En la actualidad hay más de 80 Sistemas operativos Controladores SDN construidos en distintos lenguajes [9].

Con dicho proyecto se espera alcanzar un modelo ontológico de los controladores SDN y que dicho modelo sirva para la interoperabilidad de políticas entre los controladores, como por ejemplo dos controladores que deben comunicar paquetes de video, multimediales y teleconferencia para brindar calidad de servicio a esos paquetes de dominios SDN diferentes como se muestra en la Figura 1 en la que se usó Mininet como simulador para definir el problema.

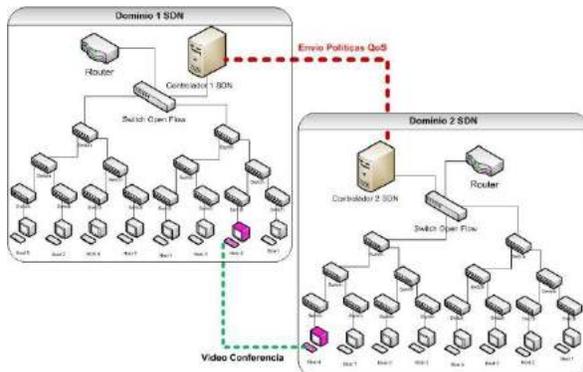


Figura 1 - Video Conferencia Envío de Política de Calidad de Servicio – Ejemplo con Mininet

2. Líneas de Investigación, Desarrollo e Innovación

Todos los modelos SDN tienen alguna versión de un controlador SDN, las API northbound las que se utilizan para comunicarse entre el

controlador SDN y los servicios y aplicaciones que se ejecutan a través de la red, pueden usarse para facilitar la innovación y permitir una automatización eficaz de la red para alinearse con las necesidades de diferentes aplicaciones a través de la programación de la red SDN según se muestra en la Figura 2.

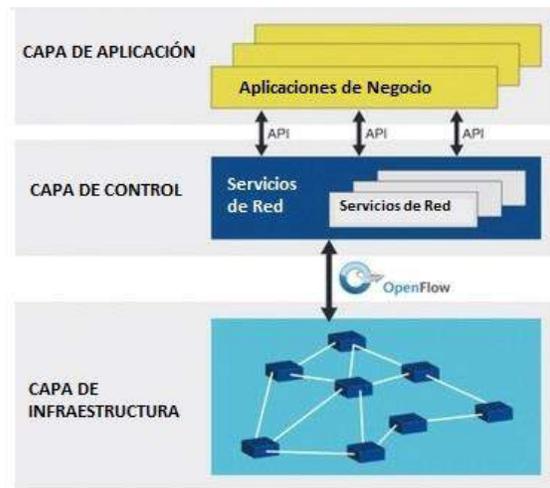


Figura 2 - Arquitectura SDN

2.1. Mapa conceptual de términos.

Como podemos ver en la Figura 3, queda reflejada el marco conceptual de términos, metodología y teoría a utilizar en esta línea de investigación.

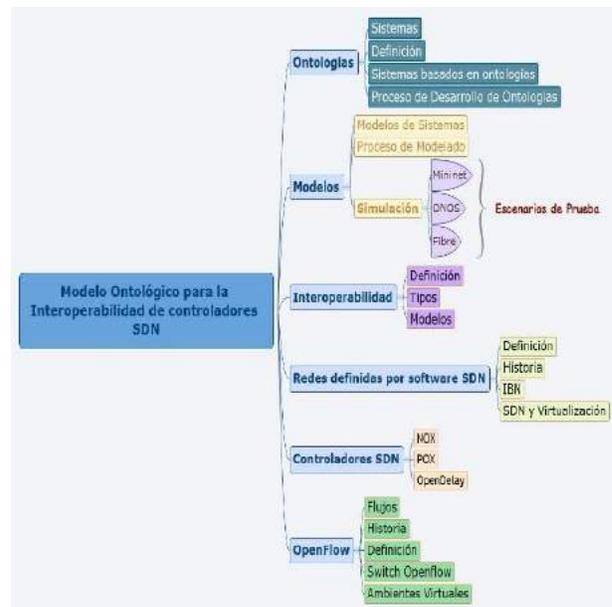


Figura 3 - Mapa conceptual de términos

Una ontología define el vocabulario de un área mediante un conjunto de términos básicos y

relaciones entre dichos términos, así como las reglas que combinan términos y relaciones que amplían las definiciones dadas en el vocabulario.

La interoperabilidad se define como la capacidad de intercambiar y compartir datos entre dos sistemas o componentes informáticos sin la intervención de un tercer sistema, de modo que la información o datos compartidos puedan ser utilizados sin requerir una comunicación previa [10].

La interoperabilidad entre sistemas no incluye solamente la habilidad de los sistemas para intercambiar información, sino también la capacidad de interacción y la ejecución de tareas conjuntas. Por tanto, el objetivo es crear un “sistema de sistemas” que no provea solamente interconectividad entre sistemas sino que logre una unión de sistemas interoperables [11].

2.2.Trabajos Relacionados

En particular podemos mencionar, Gestión de Redes de datos a través de ontologías utilizando sistemas multiagentes [12].

Este trabajo es muy similar define mediante ontologías los diferentes tipos de equipos, servidores, dispositivos que son parte de una red de computadores, para que los administradores de red puedan realizar una gestión a través de un sistema basado en ontologías, de esta forma lograr la interoperabilidad entre los diferentes dominios y elementos de gestión. [12]

Otro proyecto similar es Gestión semántica Aplicando las ontologías a la gestión de las redes. Es muy focalizado en los diferentes modelos de gestión de redes integrados que usan distintas tecnologías como SNMP, CMIP, DMI y WBEM. [13]

Por otro lado otro proyecto similar es “Un enfoque basado en la ontología hacia la configuración de dispositivos de red heterogéneos”. Una de las tareas más complejas y esenciales en la gestión de redes es la configuración de equipos. La falta de mecanismos estandarizados para la modificación y control de la configuración de equipos ha llevado al uso continuado y extendido de interfaces por líneas de comando

(CLI). Desafortunadamente, las CLIs son generalmente, específicas por fabricante y dispositivo. Se presenta el diseño y especificaciones de [14].

3. Resultados Obtenidos/Esperados

3.1.Definición del Problema

La solución propuesta a los problemas mencionados previamente consiste en la definición de un modelo ontológico que condense el vocabulario común (conceptos) y las relaciones (entre conceptos) existentes de SDN y sus controladores. Para llevar a cabo esta tarea se toma como referencia un Modelo de Interoperabilidad a Nivel Conceptual el que da rigor ingenieril a las diferentes etapas por las que pasará la definición de la Ontología. Luego de realizada la tarea mencionada previamente, se lleva a cabo una investigación profunda de los diferentes controladores en distintos escenarios simulados con herramientas como Mininet, Onos y Fibre. Esta tarea tiene como finalidad poder extraer un vocabulario común y las relaciones existentes entre los términos para poder, cumpliendo con las etapas del modelo de Interoperabilidad, dar soporte a las distintas etapas de una forma integrada, consistente y flexible [15].

3.2.Objetivos de esta investigación

Objetivo General:

Definir un modelo Ontológico para redes SDN de dominios diferentes que permita una interoperabilidad entre los controladores existentes.

Objetivos específicos son:

- Identificar datos pertenecientes a diferentes Dominio SDN y sus controladores.
- Realizar la recopilación y análisis de los datos del dominio que se investiga para desarrollar el modelo conceptual.
- Definir un modelo ontológico en dominios SDN con diferentes controladores
- Ensayar el modelo en un escenario de prueba para validarlo.

4. Formación de Recursos Humanos

La línea de investigación plantea dar conocimientos sobre Modelos en SDN y sobre dos controladores SDN pueden compartir políticas, así como también aplicar conceptos de arquitecturas en la cloud para una escalabilidad horizontal y un buen desempeño, sobre los conceptos de prioridades de comunicación en escenarios Big Data, Data Center y Videoconferencia.

Dentro del proyecto se contempla la formación de alumnos de las cátedras Redes de Información, Redes Avanzadas, Calidad de Software en los aspectos referentes a métodos de Redes definidas por software, en Modelos, Otologías y en aspectos de cómo aplicar conceptos de prioridades de comunicación en escenarios Big Data, Data Center y Videoconferencia.

En el caso de los integrantes del grupo existen 3 integrantes que están realizando su tesis final para el término de la carrera de la maestría en Ingeniería en Software, para los cuales, este proyecto será de suma ayuda y les aportará no solo metodología sino conocimientos específicos de las temáticas planteadas para cada tema de tesis.

Referencias

- [1] M. Rouse, «Searchsdn Techtargt,» 08 2015. [En línea]. Available: <http://searchsdn.techtargt.com/definition/software-defined-networking-SDN>. [Último acceso: 02 2016].
- [2] CCNA, «ccna-certification,» 25 10 2015. [En línea]. Available: <http://www.ccna-certification.info/que-es-el-software-defined-networking-sdn>. [Último acceso: 02 2016].
- [3] M. Rouse, «Definition: OpenFlow,» techtargt, 06 2012. [En línea]. Available: <http://searchsdn.techtargt.com/definition/software-defined-networking-SDN>. [Último acceso: 03 2017].
- [4] Google, Inc., «ucsd,» 2014. [En línea]. Available: <http://cseweb.ucsd.edu/~vahdat/papers/b4-sigcomm13.pdf>. [Último acceso: 02 2017].
- [5] K. Forster, «Will the Facebook switch and SDN OS change networking,» 06 2014. [En línea]. Available: <http://searchsdn.techtargt.com/opinion/Will-the-Facebook-switch-and-SDN-OS-change-networking-Maybe>.
- [6] S. M. Kerner, «Why Facebook Does SDN,» Enterprisenetworkingplanet, 03 2014. [En línea]. Available: <http://www.enterprisenetworkingplanet.com/datacenter/why-facebook-does-sdn.html>. [Último acceso: 08 2017].
- [7] RNP, «Red Nacional de Investigación de Brasil,» 2017. [En línea]. Available: <https://www.rnp.br/>. [Último acceso: 08 2017].
- [8] C. R. S. S. Natarajan, «sdnhub,» sdnhub, 2014. [En línea]. Available: <http://sdnhub.org/tutorials/openflow-1-3/>. [Último acceso: 06 2016].
- [9] S. Central, «SDN Controller Comparison Part 2: Open Source SDN Controllers,» SDX Central, 07 2016. [En línea]. Available: <https://www.sdxcentral.com/sdn/definitions/sdn-controllers/open-source-sdn-controllers/>. [Último acceso: 02 2017].
- [10] L. E. R. G. - R. d. J. G. Herrera, «SISTEMA DE INFORMACIÓN PARA APOYAR LA GESTIÓN DE LA EDUCACIÓN A DISTANCIA,» *TELEMATIQUE*, vol. 9, nº 1, p. ., 2010.
- [11] M. W. M. B. A. S. y. A. R. M.A. Manso, «Modelo de Interoperabilidad Basado en Metadatos (MIBM),» Depto. de Ingeniería Topográfica y Cartográfica. ETSI en Topografía, Geodesia y Cartografía, Universidad Politécnica de Madrid, Madrid, 2008.
- [12] W. H. CHICA, «GESTIÓN DE REDES DE DATOS A TRAVÉS DE ONTOLOGÍAS UTILIZANDO SISTEMAS MULTIAGENTES,» UNIVERSIDAD NACIONAL DE COLOMBIA SEDE MANIZALES FACULTAD DE INGENIERIA Y ARQUITECTURA MAESTRIA EN AUTOMATIZACIÓN INDUSTRIAL, Colombia, 2015.
- [13] V. V. J. B. Jorge Lopez de Vergara, «Gestión Semántica Aplicando las Ontologías a la gestión de la red,» Universidad Politécnica de Madrid, Madrid, 2015.
- [14] A. Martinez, «An Ontology-Based Approach Toward the Configuration of Heterogeneous Network Devices,» Department of Computer Architecture Technical University of Catalonia (UPC), Barcelona, 2015.
- [15] L. M. d. C. Gil, «ONTOLOGÍA PARA LA INTEROPERABILIDAD DE MODELOS DE SIMULACIÓN,» San Luis, 2015.

Figuras

- Figura 1 - Video Conferencia Envío de Política de Calidad de Servicio – Ejemplo con Mininet 3
- Figura 2 - Arquitectura SDN 3
- Figura 3 - Mapa conceptual de términos..... 3

METODOLOGÍAS Y HERRAMIENTAS EMERGENTES PARA CONTRIBUIR CON LA CALIDAD DEL SOFTWARE

Dapozo, Gladys N.; Greiner, Cristina; Irrazabal, Emanuel; Medina, Yanina; Ferraro, María; Mascheroni, Agustín, Company, Ana M., Badaracco, Numa H.
Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad
Nacional del Nordeste
{gndapozo, cgreiner, eirrazabal, mascheroni.yanina}@exa.unne.edu.ar,
mafferraro@gmail.com, anamacom@gmail.com, numahernan@gmail.com,

RESUMEN

Este proyecto es una continuación de los proyectos F07-2009 y F10-2013, ambos enfocados en modelos, métodos y herramientas para la calidad del software. En esta propuesta se abordarán temas emergentes en el área de la calidad de software, en particular, aspectos vinculados con la gestión cuantitativa de proyectos y entrega continua en entornos ágiles buscando promover y/o generar métodos y herramientas que contribuyan a mejorar la calidad del proceso y del producto software.

Una vía para la obtención de un producto software de calidad tanto en el desarrollo como en el mantenimiento es una apropiada gestión de la asignación de los pedidos de cambio al RRHH más apropiado. Si esto se combina con una medición de los atributos que permiten tener una medida de la calidad, es posible brindar a los responsables de los proyectos de desarrollo de software una poderosa herramienta para la mejora de la calidad del mismo.

Por otro lado, se está trabajando en el desarrollo de procedimientos para la gestión de las actividades de prueba continua en la disciplina de entrega continua. Esto incluye el análisis comparado de las herramientas actuales de entrega continua y la evaluación de sus características. Se continúa con métodos cuantitativos para estimar esfuerzo de desarrollo y se atiende la problemática de la eficiencia de los sistemas en organismos públicos, mediante la trazabilidad y el enfoque de procesos de negocios.

Palabras clave: calidad de software, gestión cuantitativa, entrega continua.

CONTEXTO

Las líneas de Investigación y Desarrollo presentada en este trabajo corresponden al proyecto PI-17F018 “Metodologías y herramientas emergentes para contribuir con la calidad del software”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2018-2021.

1. INTRODUCCIÓN

En un mundo globalizado y moderno, donde el software es cada vez más complejo y de gran tamaño [1], llevar adelante un proceso de desarrollo que finalice dentro de los parámetros acordados, a la vez que se entrega un producto de calidad no es una tarea trivial, así como tampoco lo es mantener esa calidad posteriormente.

La calidad del software es una compleja combinación de factores, que varían entre diferentes aplicaciones. Diversos autores como Pressman [2], McCall [3], y estándares tales como la ISO 25010 [4] han tratado de determinar y categorizar los factores que afectan a la calidad del software. Y en general se concluye que la calidad interna del producto software está directamente relacionada con diferentes atributos cuantificables a partir de su código fuente. Un ejemplo de ello son las métricas CK de diseño orientado a objeto [5], o las métricas de Halstead [6].

Asimismo el análisis del código fuente a partir de estas métricas puede ser tenido en

cuenta para describir la complejidad del producto software y la probabilidad de encontrar errores [7].

Una vez identificados los problemas en el código fuente, corregirlos no resulta una tarea trivial, sobre todo en proyectos grandes. Es esencial y deseable una gestión adecuada de los pedidos de cambio que atraviesa un proyecto, desde el momento que es reportado, asignado hasta que se considera finalizado.

El proceso de asignación de recursos humanos de la manera más eficiente posible puede tener un efecto directo en el cumplimiento de parámetros establecidos, evitando pérdidas de tiempo, esfuerzo y calidad del producto.

Encontrar el desarrollador apropiado para un cambio solicitado es una tarea fundamental para su resolución de la manera más eficiente [8]. Sin embargo, resulta una tarea difícil de automatizar y que consume tiempo. Requiere de diferentes habilidades cognitivas, comunicación con miembros del equipo y conocimiento del proyecto, tanto de los aspectos técnicos como organizacionales.

Siguiendo con este mismo enfoque y con el objetivo de lograr una mayor satisfacción del cliente, es crítico para las organizaciones entregar productos de calidad de manera aún más rápida. Esto dio lugar al surgimiento de un nuevo enfoque denominado “Entrega Continua de Software”, más conocido en inglés como Continuous Delivery (CD). En este enfoque los equipos mantienen la producción de software en ciclos cortos de tiempo, asegurando que el producto pueda ser lanzado de manera fiable en cualquier momento [9] [10].

Existe un concepto similar al CD, que es el de Despliegue Continuo, en inglés Continuous Deployment (DC). El DC es una actividad que consiste en lanzar cambios continuamente al ambiente de producción. La principal diferencia se encuentra en la fiabilidad a la hora de lanzar una nueva versión del producto: el DC busca integrar código a producción una, dos, y muchas

más veces en el mismo día, en cambio, el CD se centra en hacerlo con la certeza de que el producto que se está lanzando a producción tiene un alto grado de calidad y se encuentre libre de defectos.

Uno de los principales problemas de estos enfoques estaría en la calidad del producto software. Ésta puede disminuir, dado que, al realizarse los despliegues del sistema con mayor frecuencia, aparecen más defectos en el producto [10]. Por tanto, es esencial desarrollar un enfoque de priorización de los diferentes aspectos en la calidad del producto software, teniendo en cuenta la forma de trabajo actual de las empresas de desarrollo software.

Para adoptar este enfoque, se utiliza el concepto de Tubería de Despliegue (DP - Deployment Pipeline), un estándar para automatizar el proceso de CD [11].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación, se describen las siguientes líneas de investigación y desarrollo:

Estimación

Una estimación es “una predicción de cuánto tiempo durará o costará un proyecto”, y constituye la base para la planificación de los proyectos. Y en particular, la estimación del esfuerzo es una tarea principal en la gestión de un proyecto software. Se estudiarán los conceptos, métodos y herramientas orientados a la estimación, en particular para ambientes de desarrollo ágil, con el objetivo de desarrollar un modelo de implementación adecuado y una herramienta que lo soporte, teniendo en cuenta la gran cantidad de técnicas ágiles relacionadas con la estimación y los diferentes grados de adopción de agilísimo.

Entrega continua

Se plantea analizar los tipos de pruebas que existen en la actualidad y los enfoques más actuales para agilizar el proceso de pruebas (como comparación de imágenes o pruebas con inteligencia artificial). Una vez hecho este análisis, se procederá con la selección

de las que más se adapten a las necesidades del Despliegue Continuo.

El paso siguiente es la creación de un primer modelo formado por distintas etapas de pruebas basándose en la Tubería de Despliegue. Para validar el modelo, se procederá con el desarrollo de un marco de trabajo que implemente estas pruebas. Se hará una serie de experimentos con este marco para comparar la velocidad de ejecución de las pruebas contra las convencionales. Se propone implementar el modelo en un ambiente real de desarrollo, para analizar los resultados obtenidos. Se someterá el modelo a mejoras, a medida que avance la etapa de implementación.

Gestión de proyectos en el sector público

Los proyectos de desarrollo software en la Administración Pública requieren cada vez mayor eficiencia, respecto a la optimización en el uso de los recursos involucrados; pero además deben lograr sumar apoyo en la estructura del estado y su cultura organizacional.

Una de las técnicas que pueden ser llevadas adelante para lograr mejorar la eficiencia de los sistemas en este tipo de organizaciones es la trazabilidad de los proyectos. Su importancia se resalta si consideramos que la trazabilidad en la Ingeniería de Software es una práctica de control que ayuda a obtener el producto en el dominio de la solución lo más exacto y fiable posible a las necesidades expresadas por el cliente en el dominio del problema.

Otro mecanismo que se prevé abordar en esta línea de investigación, es la identificación y modelado de procesos de negocios para guiar el diseño de aplicaciones y mejorar la usabilidad de las mismas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados provienen en parte del proyecto de investigación F010-2013: “Métodos y herramientas para la calidad del software”, finalizado en el año 2017.

A continuación, se indican los antecedentes:

- En [12] se realizó una revisión sistemática de la literatura. En ella, se buscaron propuestas, técnicas, enfoques, métodos, herramientas y otro tipo de soluciones para afrontar los problemas de pruebas en entornos de desarrollo continuo.
- En [13] se llevó a cabo una revisión sistemática con el objetivo de identificar, analizar y sintetizar las técnicas, herramientas y desafíos encontrados en la literatura sobre pruebas de compatibilidad web. Los resultados indicaron que la técnica de prueba más elegida es el análisis visual.
- En [14] se presentó una técnica para automatizar la detección de defectos de incompatibilidad, mediante el uso del procesamiento digital de imágenes.
- En [16] se propuso un patrón de diseño para realizar pruebas automatizadas sobre servicios web de tipo RESTful, basados en los principios de la arquitectura REST y las pruebas unitarias.
- En [16] se propuso una herramienta para la gestión de proyectos de software que integra la asignación y seguimiento de pedidos de cambios con el análisis de métricas aplicadas al código fuente. La información que la misma ofrece permitirá a los responsables y participantes del proyecto tener una visión amplia de los avances y retrasos del proyecto, y a la vez, un panorama de la calidad del mantenimiento correctivo y adaptativo en función de la medición del código fuente y los valores umbrales definidos.

En cuanto a Resultados Esperados, se trabaja actualmente en:

- Propuesta de estimación de software en contextos de desarrollo ágil.
- Elaboración de un marco de trabajo para guiar el diseño de aplicaciones, basado en el modelado de procesos de negocios para mejorar la usabilidad de las

aplicaciones en el contexto de un organismo de la administración pública provincial.

4. FORMACIÓN DE RECURSOS HUMANOS

En el Grupo de Investigación sobre Calidad de Software (GICS) están involucrados 6 docentes investigadores, 1 tesista de doctorado, 1 becario de investigación de pregrado y 2 tesistas de maestría que finalizaron el cursado en el 2017. Un alumno de grado finalizó la carrera Licenciatura en Sistemas de Información con un trabajo vinculado al proyecto.

5. REFERENCIAS

- [1] M. Lehman, "On Understanding Laws, Evolution and Conservation in the Large Program Life Cycle", *J. of Sys. and Software*, vol. 13, pp. 213-221, 1980.
- [2] R. Pressman, *Ingeniería de Software. Un Enfoque Práctico*, Sexta ed., McGraw Hill, 2005.
- [3] J. A. McCall, P. K. Richards y G. F. Walters, *Factors in Software Quality*, 1997.
- [4] International Organisation for Standardisation, *ISO/IEC 25010, Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuRE) -- System and software quality models*; Geneva, Switzerland, 2011.
- [5] S. Chidamber y C. Kemerer, «A metrics suite for object oriented design,» *Software Engineering, IEEE Transactions*, vol. 20, n° 6, pp. 476 - 493, 1994.
- [6] H. H. Halstead, *Elements of Software Science*, Nueva York: Elsevier Science Inc., 1977.
- [7] R. Malhotra y A. Jain, «Software Fault prediction for Object Oriented Systems: A Literature Review,» *ACM SIGSOFT Software Engineering Notes*, vol. 36, n° 5, pp. 1-6, 2011.
- [8] Y. C. Cavalcanti, P. A. da Mota Silveira Neto y J. do Carmo Machado, «Towards Understanding Software Change Request Assignment: a survey with practitioners,» de *Proceedings of the 17th International Conference on Evaluation and Assessment in Software Engineering*, Porto de Galinhas, Brazil, 2013
- [9] L. Chen, "Continuous Delivery: Huge Benefits, but Challenges Too" in *IEEE Software* 03/2015. V. 32(2).
- [10] J. Humble and D. Farley. "Continuous delivery: reliable software releases through build, test, and deployment automation", 1st ed. Boston, US: Pearson Education, 2010.
- [11] O. Prusak. "Continuous Testing: The Missing Link in the Continuous Delivery Process". *Blaze Meter*. 2015. <https://blazemeter.com/blog/continuous-testing-missing-link-continuous-delivery-process>
- [12] Mascheroni, M.A., Irrazábal, E. (in press). Continuous Testing and solutions for testing problems in Continuous Delivery: A Systematic Literature Review. *Computación y Sistemas*.
- [13] Sabaren, L; Mascheroni, M.A.; Greiner, C; Irrazábal, E. Una Revisión Sistemática de la Literatura en Pruebas de Compatibilidad Web XXIII Congreso Argentino de Ciencias de la Computación CACIC 2017. ISBN: 978-987-4251-43-5. La Plata (Argentina). Fecha: Octubre de 2017
- [14] Mascheroni, M.A., Irrazábal E, Cogolio M.K., Automatic detection of Web Incompatibilities using Digital Image Processing. *SADIO Electronic Journal of Informatic and Operation Research*. EJS 16 (1) 29-45 (2017)29
- [15] Mascheroni, M.A., Irrazábal, E. (in press). A Design Pattern Approach for RESTful tests: A case study. *IEEE 12th Colombian Computing Congress*.
- [16] Chiapello, Jorge. "Gestión de Pedidos de Cambio y Métricas Orientadas a Objetos". Proyecto Final de Carrera.

Licenciatura en Sistemas de
Información. UNNE. Defendido el 27-
12-2017.

Modelo de interoperabilidad entre sistemas para trazabilidad de procesos internos de la industria cárnica empleando nuevas tecnologías de identificación

Ing. Paez Sergio, Ing. Calloni Juan, Ing. Bianciotti Andrés,
Banchio Leandro, Scharff Lucía

Grupo PowerTICs / Secretaría de Ciencia y Tecnología
Facultad Regional San Francisco / Universidad Tecnológica Nacional
Av. de la Universidad 501, 03564-421147
sergio_paez@hotmail.com, andresbianciotti@gmail.com

RESUMEN

Este trabajo presenta un modelo de estandarización e inclusión de nuevas tecnologías de identificación utilizadas en la industria cárnica para crear una base de conocimientos sobre el proceso productivo (trazabilidad interna) basado en el sistema de trazabilidad exigido por los organismos reguladores, como ser SENASA.

La base de conocimiento permitirá lograr la interoperabilidad de los distintos sistemas existentes en el proceso productivo como ser MRP (planificación de los requerimientos de materiales), PCP (planificación y control de la producción), CRP (planificación de requerimientos de capacidad), Gestión de Stock, Control de Calidad, HACCP, etc

Palabras clave: Trazabilidad, base de conocimiento, interoperabilidad, procesos productivos, RFID (Radio Frequency Identification), WSN (redes de sensores inalámbricos), IoT (Internet of Things).

CONTEXTO

La línea de investigación presentada se encuentra en el marco de las áreas prioritarias para el desarrollo de las actividades de I+D que se formalizaron a través de la Resolución de Consejo Directivo N 353/2016 de la Universidad Tecnológica Nacional Facultad Regional San Francisco cuyas áreas prioritarias son: gestión de procesos de negocios, ingeniería de software, gestión y tecnologías de las organizaciones, calidad de

Software, seguridad de la información y bases de datos.

Se involucran diversas áreas de estudio como la Ingeniería de Software relacionándose con el desarrollo de sistemas a partir de arquitecturas hexagonales, las Bases de Datos mediante la normalización y persistencia unificada de datos en una base de conocimiento, la Gestión de Procesos de Negocios (BPMN) y dispositivos y tecnologías de comunicación de vanguardia (RFID, IoT y WSN). Estas líneas se encuentran insertas en el Proyecto de Investigación y Desarrollo (PID) N° EIUTNSF0003885 “Desarrollo de un framework para la interoperabilidad de sistemas de trazabilidad mediante modelo ontológico semi formal”, el mismo se encuentra en desarrollo desde el mes de Mayo de 2016. Dicho proyecto es ejecutado por el grupo de I+D PowerTICs de la UTN Facultad Regional San Francisco.

El proyecto se encuentra homologado y financiado por la Secretaría de Ciencia y Tecnología de la Universidad Tecnológica Nacional.

1 - INTRODUCCIÓN

En la industria alimenticia es necesario contar con un sistema de trazabilidad que permita la mejora de los controles de calidad, la optimización de los procesos productivos, el aumento de la confianza del consumidor y en casos de ser necesario el retiro del mercado de partidas específicas de productos. La obligación de su implementación por parte de las empresas radica en cumplimentar con la

legislación vigente impuesta por organismos de control públicos como ser para en el caso de la industria de la carne el SENASA. Las experiencias y los informes del desarrollo de aplicaciones de trazabilidad muestran que la trazabilidad es tarea costosa y complicada [1].

Los sistemas de trazabilidad interna mantienen registros para un determinado proceso de los lotes de entrada y su correlación con los lotes de salida a través de las operaciones que sufre el lote durante el proceso [2], tal como se muestra en la Figura 1.

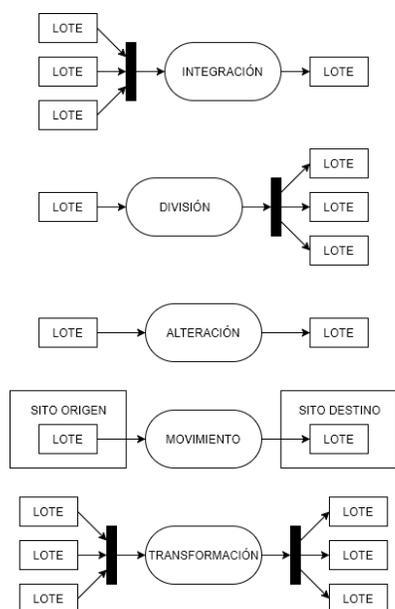


Figura 1 – Operaciones básicas sobre lotes de producción.

Este trabajo propone el uso de nuevas tecnologías de identificación para automatizar el registro en los puntos de control del sistema de trazabilidad y asegurar la validez y exactitud de los datos medidos [3]. A partir de estos registros se genera una base de conocimiento con los datos de producción sobre cada lote que da soporte a las necesidades de información de los distintos sistemas.

1.1. Conceptos de trazabilidad

Según la International Organization for Standardization [4], la trazabilidad se relaciona con la capacidad para seguir

históricamente una aplicación o localización de algo que esté bajo consideración u observación. Esta definición se centra en el rastreo de un producto o servicio en un horizonte o línea de tiempo, sin tener en cuenta el sistema de trazabilidad particular.

Según el Codex Alimentarius, “Trazabilidad es la capacidad para seguir el movimiento de un alimento a través de etapa(s) especificada(s) de la producción, transformación y distribución” [6].

1.2. Tecnologías de identificación actuales

El relevamiento realizado en una empresa tomada como caso de estudio, arrojó las siguientes tecnologías empleadas para identificación de lotes, mezclas, productos intermedios y productos terminados: códigos de barras y códigos QR impresos sobre etiquetas, inscripciones realizadas sobre el envase primario y/o secundario del producto. Estas tecnologías proporcionan datos reducidos a cada etapa del proceso productivo debido a la limitación de caracteres que pueden representar y/o la necesidad de contar con lectores y dispositivos de impresión específicos.

1.3. Nuevas tecnologías de identificación

Una red de sensores inalámbricos (WSN) consiste en nodos de sensores con capacidad de captura, procesamiento y transmisión de datos (Figura 2). Los nodos de sensores capturan información del ambiente, tales como temperatura y humedad, la transforman a formato digital y la transmiten de manera autónoma mediante interfaces de comunicación hacia una estación base en donde es almacenada [3].



Figura 2 – Lecturas de un sensor de temperatura y humedad empleando WSN [3].

RFID es una tecnología que transmite la identidad, en forma de número de serie único, de un objeto de manera inalámbrica, empleando ondas de radio (Figura 3), sin requerir la intervención humana [3].

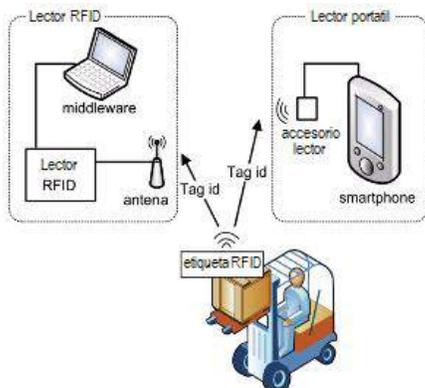


Figura 3 – Arquitectura de un sistema de identificación autónoma empleando etiquetas y lectores RFID [3].

El uso de etiquetas RFID en el sistema de trazabilidad proporciona eficiencia en comparación con los códigos de barras y las etiquetas de papel. Los sistemas de trazabilidad basados en RFID se han implementado en muchas áreas, tales como la cadena de suministro de la acuicultura [7], productos lácteos [8], hortalizas frescas [9] y animales o carne vacuna [10]. La integración de RFID y WSN proporciona un monitoreo continuo de los datos a lo largo del proceso productivo y de la cadena de suministro de alimentos.

1.4. Sistemas relacionados con procesos productivos

MRP (planificación de los requerimientos de materiales), PCP (planificación y control de la producción), CRP (planificación de requerimientos de capacidad), Sistema de Gestión de Stock, Sistema de control de calidad (QMS), Análisis de Peligros y Puntos Críticos de Control (APPCC o HACCP, por sus siglas en inglés).

1.5. Trazabilidad de proceso (interna)

Tal como lo establece SENASA en el documento Bases para la implementación de un sistema de trazabilidad, en la trazabilidad interna se trata de relacionar los productos que se han recibido en la empresa, las operaciones o procesos que éstos han seguido (equipos, líneas, cámaras, mezclado, división, etc.) dentro de la misma y los productos finales que salen de ella [6].

2 - LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

El desarrollo e investigación se basa en encontrar técnicas de comunicación entre los diferentes sistemas que controlan los procesos productivos de las empresas manufactureras con el fin de brindar interoperabilidad y unificar todos los datos recabados en una misma base de conocimiento central. Se incluyen además tecnologías de vanguardia en cuanto a identificación y seguimiento autónomo de información con el fin de evitar tareas de tipo manual en la carga de información, contando con datos certeros y oportunos durante todo el proceso productivo que doten de puntos de control más específicos a los encargados del seguimiento de la trazabilidad y a los entes reguladores.

3 – RESULTADOS OBTENIDOS/ESPERADOS

Este trabajo presenta un modelo de sistema para cumplir con la trazabilidad interna, mejorando las tecnologías de identificación y

captura de datos y unificando toda la información en un repositorio compartido por los demás sistemas de la empresa.

Se propone la inclusión de nuevos puntos de control empleando WSN y RFID dentro del proceso productivo que junto con los existentes unifiquen sus registros para un mejor seguimiento del proceso (Figura 4).

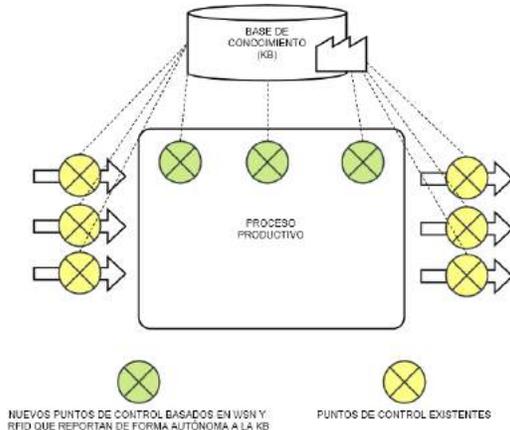


Figura 4 – Nuevos puntos de control dentro del proceso productivo que complementan los existentes.

Se creará una base de conocimiento unificada que centralice los datos de los puntos de control y sirva como repositorio común a los demás sistemas relacionados con los procesos productivos (Figura 5).

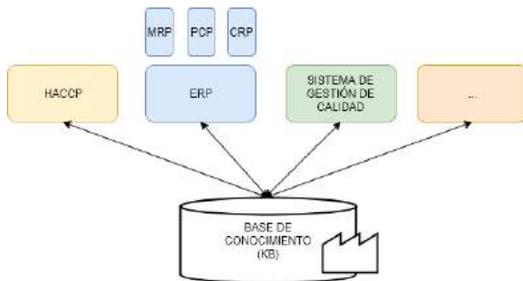


Figura 5 – Base de conocimiento de trazabilidad de procesos actuando como repositorio unificado de los demás sistemas.

Para persistir la base de conocimiento consideramos factible emplear el modelo de trazabilidad propuesto por Bechini [2].

El nuevo modelo mejora los procesos productivos al automatizar la captura de datos

de trazabilidad en los diferentes puntos de control. A su vez, introduce información de espacio-tiempo mediante los registros de sensores inalámbricos ubicados estratégicamente, lo cual permite analizar el proceso considerando nuevas dimensiones y perspectivas. Además brinda interoperabilidad a los distintos sistemas existentes en la empresa (control de calidad, HACCP, MRP, CRP, gestión de stock, etc.) tomando como componente central la base de conocimiento de trazabilidad (Figura 6).

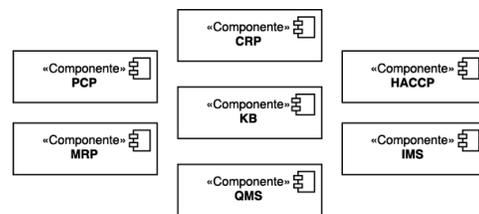


Figura 6 – Relación de la base de conocimiento (KB) con los demás sistemas industriales.

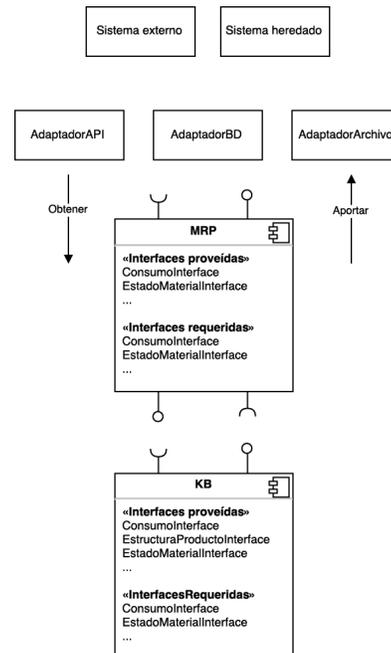


Figura 7 – Interfaces y adaptadores propuestos entre la base de conocimiento y el sistema MRP.

El modelo propone una doble direccionalidad entre los sistemas heredados o externos y la base de conocimiento (KB). Adoptando una arquitectura «ports and adapters» [11], también conocida en

implementaciones particulares como «arquitectura hexagonal», éste enfoque permite separar la lógica del dominio en un sólo punto aportándole flexibilidad a la hora de adaptarlo a reglas de negocio cambiantes y específicas (Figura 7).

La base de conocimiento (KB) es el componente central. En éste las entidades de negocio concentran la información proveniente del resto de las fuentes que luego es vinculada internamente permitiendo trazar estados y características de los productos.

Dependiendo de las interfaces ofrecidas por los sistemas externos, es posible realimentar a éstos con información resultante de la base de conocimiento y mejorar los procesos internos de cada uno.

4 - FORMACIÓN DE RECURSOS HUMANOS

El proyecto incorpora alumnos y graduados de la carrera Ingeniería en Sistemas de Información de la UTN Facultad Regional San Francisco en conjunto con estudiantes de Maestría en Ingeniería de Software y Maestría en Calidad de Software de la Universidad Nacional de San Luis (UNSL). El proceso de investigación permitirá que los participantes tengan contacto con tecnologías y metodologías novedosas, poco frecuentes y candidatas a masificarse en los próximos años: tecnologías de comunicación, tecnologías de identificación, tecnologías de desarrollo, arquitecturas y patrones de software, sistemas de gestión industrial y ontologías.

Algunos de los temas abordados dependen de conocimientos que son ajenos a la profesión de los involucrados. Esto dará lugar a que se planifiquen entrevistas con especialistas en cada dominio al que apunta la

solución fomentando la capacitación de los integrantes mediante el contacto con el medio.

5 - BIBLIOGRAFÍA

[1] Bai Qu, Xinchao Jing, Xiaojun Wang, Y.L. and Y.L., 2012. "Design on Cucumber Traceability System Based on the Internet of Things". In Daoliang Li & Y. Chen, eds. *Computer and Computing Technologies in Agriculture V*, IFIP International Federation for Information Processing 2012. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 199-208.

[2] Alessio Bechini, Mario G.C.A. Cimino, Francesco Marcelloni*, Andrea Tomasi, "Patterns and technologies for enabling supply chain traceability through collaborative e-business ", *ScienceDirect, Information and Software Technology 50* (2008) 342–359.

[3] Ganjar Alfiana*, Jongtae Rhee, Hyejung Ahn, Jaeho Lee, Umar Farooq, Muhammad Fazal Ijazb, M. Alex Syaekhoni, "Integration of RFID, Wireless Sensor Networks, and Data Mining in an e-Pedigree Food Traceability System ", *Journal of Food Engineering Volume 212*, November 2017, Pages 65-75

[4] International Organization for Standardization-ISO, "Norma Internacional ISO 9000, Sistema de gestión de la calidad ", *Fundamentos y vocabulario*. Ginebra, 2005.

[6] "Bases para la implementación de un sistema de trazabilidad", Enero 2010. SENASA - Dirección Nacional de Fiscalización Agroalimentaria - Dirección de fiscalización de productos de origen animal.

[7] Alfredo Parreño-Marchante, Alejandro Alvarez-Melcona, Mira Trebar, Piero Filippin, "Advanced traceability system in aquaculture supply chain", *Journal of Food Engineering*, Volume 122, February 2014, Pages 99-109

[8] P. Barge, P. Gay, V. Merlino, C. Tortia, "Item-level Radio-Frequency Identification for the traceability of food products: Application on a dairy product", *Journal of Food Engineering Volume 125*, March 2014, Pages 119-130.

[9] Mainetto, L., Francesca, M. Patrono, L., Simone, F., Laura, M.S., Vergallo, R., "An RFID-based tracing and tracking system for the fresh vegetables supply chain", *Int. J. Antennas Propag.* 2013, 15.

[10] Jianying Feng, Zetian Fu, Zaiqiong Wang, Mark Xu, Xiaoshuan Zhang, "Development and evaluation on a RFID-based traceability system for cattle/beef quality safety in China", *Food Control Volume 31*, Issue 2, June 2013, Pages 314-325.

[11] Hexagonal architecture: Alistair Cockburn, online at <http://alistair.cockburn.us/Hexagonal+architecture>

Línea de Investigación en Evaluación de Productos *Software*

Paula Angeleri, Rolando Titiosky, Jorge Ceballos, Sergio Aguilera

Facultad de Ingeniería y Tecnología Informática

Universidad de Belgrano

Av. F. Lacroze 1947, CABA

{paula.angeleri; rolando.titiosky; jorge.ceballos; sergio.aguilera}@comunidad.ub.edu.ar

RESUMEN

El objetivo de este artículo es presentar la situación actual y los avances realizados en el proyecto de investigación MyFEPS *Metodologías y Framework para la Evaluación de Productos de Software* [1]. MyFEPS inició como un proyecto de investigación en 2010, pero ya cumplidos sus objetivos iniciales, continuó en forma ininterrumpida, transformándose en una línea de investigación que proporciona continuidad, articulación y productividad en la sub área de Evaluación de productos software y tecnologías, manteniendo el dinamismo que promueve Tamayo [2] en la Facultad de Ingeniería y Tecnología Informática de la Universidad de Belgrano (UB). En este contexto se describen algunas de las actividades más recientes, y sus objetivos específicos.

Palabras clave: Calidad de software, Evaluación de calidad de producto *software*, Framework MyFEPS, Modelo de calidad de producto software QSAT.

CONTEXTO

MyFEPS nace en UB para dar apoyo al IRAM¹ en su servicio de certificación de productos software [3], con el objetivo de especificar un modelo de calidad actualizado [4], y un proceso de evaluación que permitiese la ponderación de características de calidad en base a objetivos de evaluación [5], entre otros

¹ Instituto Argentino de Normalización y Certificación.

objetivos explicados anteriormente [6][7][8][9][10].

1. INTRODUCCIÓN

MyFEPS continua su ciclo de vida con nuevas y variadas pruebas de campo que permiten validar, ajustar y ampliar el método de evaluación [5], el modelo de calidad QSAT [4], y la base de herramientas de apoyo al proceso de evaluación que proporciona el Framework MyFEPS [11].

Avances alcanzados

A continuación, se presentan los últimos avances de este proyecto de investigación que, con los años, ha madurado esta nueva línea de investigación de UB, la línea de investigación en evaluación de productos *software*. Se resumen los principales trabajos, y en esta publicación se hará foco en los proyectos JuegosFPS² [12] y BckDisk [13],

Evaluación de motores para video juegos FPS

El objetivo principal del trabajo final de carrera (TFC) del tesista Pablo Malerba [12] fue la comparación de motores para el desarrollo de videojuegos del tipo FPS, buscando comparar sus características más importantes, para lograr identificar entre las alternativas del mercado evaluadas, cual se adecúa mejor para un proyecto de desarrollo de videojuegos FPS, y qué ventajas proporciona.

² FPS por su sigla en inglés First Person Shooter.

Se estudiaron puntualmente tres *Game Engines*: Unity 3D, Unreal Engine y Cry Engine.

El proceso de evaluación de la calidad utilizado fue MyFEPs, con el que se definieron las etapas, las estimaciones de costo, los tiempos necesarios y los recursos humanos y sus roles del proyecto.

El modelo de Calidad elegido fue el modelo QSAT al que, durante el proyecto, hubo que hacerle ajustes necesarios para evaluar motores de juegos. Para llevar a cabo el proyecto, se evaluaron las características básicas de QSAT: CB1 “Calidad de los Artefactos” (el tesista la llamó Desarrollo), CB2 “Efectividad”, CB3 “Facilidad de mantenimiento” (el tesista la llamó Mantenibilidad), CB4 “Satisfacción de los usuarios”, CB5 “Usabilidad Objetiva”. Las subcaracterísticas evaluadas fueron CB1 SC1 “Cantidad de líneas de código”, SC2 “tiempo de desarrollo”, CB2 SC1 “Ausencia de funcionalidades”, CB2 SC2 “Cobertura de funcionalidades”, CB3 SC1 “Portabilidad”, CB4 SC1 “Jugabilidad”, CB5 SC1 “En el acceso a las funciones”.

Los resultados obtenidos durante la ejecución de los casos de prueba fueron ponderados y se obtuvo la valoración final que da sustento a la recomendación de uno de los motores evaluados. Esta recomendación no se incluye para no desfavorecer a los otros productos del mercado.

Evaluación de BckDisk

Este proyecto TFC del tesista Facundo Favaloro [13] evaluó el tipo de sistema de Backup Empresarial, comparando el sistema “Avamar + Data Domain” de la empresa EMC, y TSM de IBM, concluyendo las ventajas y desventajas de utilizar un método u otro.

El proceso de evaluación de la calidad utilizado fue MyFEPS. El modelo de calidad elegido fue el QSAT. Las características básicas evaluadas fueron “Efectividad”, “Eficiencia”, “Satisfacción de los Stakeholders que no son usuarios”,

“Satisfacción de los usuarios”, y “Seguridad Informática”. Las Métricas utilizadas fueron tomadas, en casi todos los casos, del modelo QSAT.

Se llegó a una conclusión sobre las tecnologías comparadas que no se incluye por las razones mencionadas.

Para más información, consultar las tesis de de la Facultad de Ingeniería y Tecnología Informática de la Universidad de Belgrano citadas.

2 LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas principales de investigación sobre las que se ha trabajado desde el inicio son: el Dominio de conceptos (se ha definido un Glosario de términos relacionados a la calidad del software), el Modelo de calidad de producto software (se ha definido el modelo QSAT) [4], el Meta modelo y Método de Evaluación [5], y las Herramientas de apoyo a la evaluación [11].

Respecto del proceso de evaluación que es la base del Método de evaluación MyFEPS [5], basado en la serie de normas ISO/IEC 14598 [14] y en la revisión ISO/IEC 25040 [15], se han investigado los Casos de uso más importantes que influyen en el proceso de evaluación, como ser la Definición de los objetivos de la evaluación considerando las partes interesadas, la Determinación del modelo de calidad, las características, sub características, sub-sub características, atributos y métricas a utilizar durante la evaluación, la Elaboración del Plan de evaluación, el Diseño de los casos de prueba, la Ejecución de las pruebas y la Ponderación e Integración de Resultados.

Una sub línea de investigación estudió la Intercambiabilidad del Modelo de calidad QSAT respecto de los Modelos de calidad ISO/IEC 9126 [16] e, ISO/IEC 25010 [17],

a lo largo de diferentes tesis y trabajos académicos.

En el TFC de Diego Ardizzone se mejoró el formato del Informe de resultados de la evaluación en un desarrollo BI que diera más significado a la Alta Dirección [18].

La sub línea de investigación, abordada por los tesis Santiago Christello [19], Ezequiel Romanin [20], y Christian Álvarez Worobey [21] estudió la evaluación de la calidad en otros contextos, como la planificación de un proyecto de emprendimiento de software, la calidad interna durante el diseño del producto, y la calidad externa durante el desarrollo, respectivamente. Se espera contar con más casos de estudio.

La sub línea de investigación abordada por el tesis Ventura [22] evaluó la facilidad de la primera instalación, y de reinstalación del software, entre otras características.

Otra sub línea de investigación, la iniciaron los tesis Santi [23][24] y Martínez [25] [26], que basaron sus TFC en la evaluación de la “usabilidad”, con foco en la “usabilidad objetiva”.

La sub línea de investigación en evaluación de juegos de computadora la inició el tesis Pablo Malerba con la comparación de motores para videojuegos durante el proyecto FPS arriba explicado [12].

Otra sub línea de investigación arriba mencionada es la comparación de Tecnologías que involucran software, proyecto BckDisk [13]. Estas dos experiencias abrieron dos sub líneas de investigación a desarrollar más profundamente: la de evaluación comparativa de software o de tecnologías de propósito similar, y la de evaluación ágil de software. Se espera continuar con estas dos sub líneas en 2018 y 2019.

Durante el proyecto MyFEPS se ha validado que el método de evaluación MyFEPS es adecuado para realizar evaluaciones de calidad de producto software (incluso por estudiantes no expertos), y se han probado como válidas

algunas métricas del modelo QSAT, como ser aquellas que aplican a la evaluación de las características básicas “Calidad de los artefactos”, “Efectividad”, “Facilidad de mantenimiento”, “Manejo de fallas”, “Satisfacción subjetiva de los usuarios”, “satisfacción de stakeholders no usuarios”, Usabilidad Objetiva”, “Facilidad de Instalación”, a modo de ejemplo. Pero se considera apropiado continuar esta línea de investigación MyFEPS con evaluaciones que permitan validar la adecuación de las restantes características del modelo QSAT [4].

3 RESULTADOS OBTENIDOS

El objetivo principal de los últimos años de MyFEPS fue la continua experimentación y uso del Framework MyFEPS (tanto con el modelo de calidad ISO/IEC 25010 como con QSAT) en diferentes tipos de proyectos académicos, investigación, tesis de grado, transferencia académica y transferencia en la industria, por ej. en [10][11], buscando validar y mejorar el Framework MyFEPS.

Como Objetivos alcanzados podemos mencionar, en forma no exhaustiva:

1. Se continuó validando la adecuación del modelo QSAT y del proceso de evaluación, y se mejoraron las herramientas de apoyo de MyFEPS, con el aporte del becario Eduardo Tarek Chabeldin, y del tesis Facundo Scipioni [27].
2. Se comprobó que MyFEPS es útil en evaluaciones de calidad SQuaRE (serie ISO/IEC 25000 [28]), en proyectos tales como eCom[23][24] y Morpho [10][29].
3. Se mantuvo actualizado el acervo de documentos correspondientes a métodos-logías, técnicas, modelo, y herramientas de MyFEPS.
4. Se utilizó MyFEPS en diferentes escenarios, tanto metodológicos (para evaluar su adaptabilidad), como del tipo evaluación y del tipo de producto evaluado (en desarrollo, en producción).
5. Se utilizó MyFEPS considerando particularidades

especiales de software de diferentes ámbitos, como ser: productos e-commerce, videojuegos, software académico, de la industria de la Salud, entre otros. 6. Se utilizó MyFEPS con diferentes stake-holders con distintos objetivos de evaluación. 7. Se utilizó MyFEPS para validar y ampliar diferentes tipos de herramientas de apoyo a las actividades de evaluación. 8. Se transfirieron los conocimientos adquiridos a la academia UB, UNNE, UBA, UP, UNSL, UNLP, UNLaM, Univ. La Punta, Univ. Morón, etc. (Argentina), en PUCP, UAP, Univ. Tacna, Univ. Chiclayo, Univ. Arequipa (Perú), en UDE (Uruguay), en Univ. Salesiana de Roma (Italia), y en UTS (Australia), a través de presentaciones, conferencias, videoconferencias, seminarios, ponencias en congresos, publicaciones, cursos, y trabajos en cátedras. 9. Se transfirieron conocimientos adquiridos a la industria, como ser a IRAM, TSOFT, DGRC, SyK, Soffa, Personal, LTSL, y empresas donde trabajaban los Tesistas (Argentina), a UNIT (Uruguay), a ACKLIS (Perú), a través de conferencias, publicaciones, cursos de capacitación, y servicios de asistencia técnica. 10. Se capitalizaron los conocimientos adquiridos, incorporándolos en el dictado de asignaturas de las carreras de pre grado, grado, en cursos de postgrado y de actualización profesional, en UB y UNLaM.

Los proyectos mencionados han cumplido sus objetivos y se han llevado a cabo conforme a lo planificado (con algún ajuste en el plazo inicial de algunos de ellos, debido a situaciones personales de tesistas). Respecto de los nuevos proyectos BckDisk y JuegosFPS también ambos cumplieron sus objetivos, y han realizado aportes al Framework MyFEPS, como ser ajustes al método MyFEPS o al Modelo QSAT.

4. FORMACIÓN DE RECURSOS HUMANOS

Directora: Mg. Paula Angeleri, 2010-actual. Co Director Dr. Amos Sorgen hasta 2012.

Profesores investigadores: Mg. Rolando Titiosky (actual), Lic. Jaquelina Wuille Bille (hasta 2013, luego extensión), Esp. Jorge Ceballos (actual), y empleados de TSOFT.

Profesores invitados que han colaborado: Lic. Alejandro Oliveros de UB (2014-2017), Mg. Sergio Aguilera (2016-actual) y Dr. Abraham Dávila de PUCP (2015-actual).

Tesistas de grado: Agustín Ventura, Martín Santi, Diego Ardizzone, Facundo Scipione, Bárbara Martínez, Ezequiel Romanin, Santiago Christello, Christian Álvarez Worobey (finalizadas), Pablo Malerba, Facundo Favaloro, Gastón Oneto, Mariano Pereda, Mauro Otonello (en curso).

Los integrantes de este grupo de investigación dirigen a estudiantes en diferentes contextos (becarios, tesinas, obligaciones académicas, etc. Además, han participado numerosos alumnos de UB y de UNLaM durante el dictado de asignaturas de grado, pregrado, y cursos (Ej. Curso de posgrado en Calidad de Software UB-INTI, cursos Empleartec), en los cuales se vienen enseñando las normas ISO/IEC de Evaluación de productos Software y el Framework MyFEPS y se han propuesto numerosos ejercicios prácticos que han aportado experiencias y puntos de mejora en el modelo de calidad de productos software QSAT, y en el framework y método de evaluación de productos software MyFEPS.

5. BIBLIOGRAFIA

- [1] Sorgen, A., Angeleri, P.: El Modelo de Evaluación del Proyecto MyFEPS, 40JAIIO -ASSE 2011 - ISSN: 1850-2792, 180-191, 2011,
- [2] Tamayo, M.: El Proceso de la Investigación Científica, Editorial Limusa, México, 2002.

- [3] Angeleri, P.: Premios Sadosky 2008, IRAM Newsletter, Año 2 Nro.6, 2009.
- [4] Angeleri, P.; Oliveros, A; Sorgen, A.; Titiosky, R.; Wuille Bille J.: Modelo de calidad de productos de software, CONAISI 2014, ISSN: 2346-9927 - Página 1043.
- [5] Angeleri, P.; Titiosky, R.; Santi, M.; Davila, A.: Proceso de Evaluación de Productos *Software*, CONAISI 2015.
- [6] Sorgen, A., Angeleri, P.: Teoría y Práctica en la Evaluación de Productos de Software, XIV Workshop de Investigadores en Ciencias de la Computación 2012.
- [7] Titiosky, R., Angeleri, P., Sorgen, A., Wuille Bille, J.: "Proyectos de Evaluación de Productos de Software con un nuevo Framework de Calidad", XV Workshop de Investigadores en Ciencias de la Computación 2013, ISBN: 9789872817961.
- [8] Angeleri, P., Titiosky, R., Sorgen, A., Wuille Bille, J., Oliveros, A.: Ajustes al Framework de Evaluación de Productos de Software MyFEPS, XVI Workshop de Investigadores en Ciencias de la Computación 2014.
- [9] Angeleri, P., Titiosky, R., Ceballos J.: Framework de Evaluación de Productos Software, XVIII Workshop de Investigadores en Ciencias de la Computación (WICC) 2016, Entre Ríos, Argentina, ISBN: 978-950-698-377-2, p. 657-661.
- [10] Angeleri, P., Titiosky, R., Ceballos J. Maspero, Ch., Sánchez A., Menal M., Vinjoy M., SOFTWARE PRODUCT CERTIFICATION: AN ACADEMIC-INDUSTRY CERTIFICATION PROJEC, 2016 IEEE Congreso Argentino de Ciencias de la Informática y Desarrollos de Investigación (CACIDI 2016), Buenos Aires, Argentina, p.176. IEEE Catalog Number: CFP16H47-POD, ISBN (Print-On-Demand): 978-1-5090-2939-6, ISBN (Online): 978-1-5090-2938-9, p. 176.
- [11] Angeleri, P., Sorgen, A., Bidone, P., Fava, A., Grasso, W.; Diseño y desarrollo de un framework metodológico e instrumental para asistir a la evaluación de software, 43 JAIIO - JULI 2014 - ISSN: 1851-2518 - Página 10.
- [12] Malerba P., Comparación de herramientas para desarrollo de videojuegos FPS, UB, 2018.
- [13] Favaloro F., Evaluación comparativa de tecnologías Backup a Disco vs Backup a Cinta, usando el Framework MyFEPS, UB, 2018.
- [14] serie de normas IRAM-ISO-IEC 14598: Tecnología de la información. Ingeniería de software. Evaluación del producto de software. Information technology. Software engineering. Software product evaluation. Parte 1, 2006, Partes 2,3, 2007, Partes 4,5, 2009, Parte 6, 2012.
- [15] ISO/IEC 25040 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Evaluation process. 2012
- [16] IRAM-NM-ISO/IEC 9126-1 Tecnología de la información-Ingeniería de software - Calidad del producto Parte 1 - Modelo de calidad (ISO/IEC 9126-1:2001, IDT) Information technology - Software engineering - Product quality Part 1 - Quality Model 1, IRAM, 2009.
- [17] ISO/IEC 25010 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality model.
- [18] Ardizzone, D., Desarrollo de una solución de *Business Intelligence* que informe el grado de calidad de un software evaluado con el Framework MyFEPS, Universidad de Belgrano, 2015. <https://biblioteca.ub.edu.ar/cgi-bin/koha/opac-detail.pl?biblionumber=50346>, 08/03/2018 09:00 hs.
- [19] Christello, S., Desarrollo Empresarial del Sistema de Información Medicinal-SIM, tesis de grado de Universidad de Belgrano, 2016. <https://biblioteca.ub.edu.ar/cgi-bin/koha/opac-detail.pl?biblionumber=51078>, 08/03/2018 09:05 hs.
- [20] Romanin, E., Análisis y Diseño del Sistema de Información Medicinal SIM, Univ. Belgrano, 2016. https://biblioteca.ub.edu.ar/cgi-bin/koha/opac-detail.pl?biblionumber=51358&shelfbrowse_itemnumber=79131#shelfbrowser, 08/03/2018 09:00 hs.
- [21] Álvarez Worobey, Ch., Rediseño de la arquitectura de una aplicación web farmacéutica, utilizando Struts 2, Universidad de Belgrano, 2017. <https://biblioteca.ub.edu.ar/cgi-bin/koha/opac-detail.pl?biblionumber=87104>, 18/03/2018 20:15 hs.
- [22] Ventura, A., Aplicación del Proyecto MyFEPS al Sistema de Gestión de Clínicas (SGC) para la Evaluación de Calidad, tesis de grado de Universidad de Belgrano, 2013.
- [23] Santi, M., Medición de la Usabilidad de un Ecommerce empleando el Framework MyFEPS, tesis de grado de Universidad de Belgrano, 2014. <https://biblioteca.ub.edu.ar/cgi-bin/koha/opac-detail.pl?biblionumber=49786>, 08/03/2018 14:50 hs.
- [24] Santi, M., Evaluación de la Usabilidad de un sitio de Comercio Electrónico empleando el Framework MyFEPS, CONAISI 2014.
- [25] Martínez, B., Algoritmo para medir el cumplimiento del atributo 17.2.1 del Modelo de calidad de productos software QSAT, U. Belgrano, 2016. https://biblioteca.ub.edu.ar/cgi-bin/koha/opac-detail.pl?biblionumber=51359&shelfbrowse_itemnumber=79133#shelfbrowser, 16/03/2018 11:30 hs.
- [26] Martínez, B., Algoritmo para Medir el Cumplimiento del Atributo 'Número de Reglas de Interfaz Estándares Cumplidas por el Sistema' CONAISI, 2017, p. 458/459.

- [27] Scipioni, F., MYFEPS, construcción escalable, U. Belgrano, 2017, <https://biblioteca.ub.edu.ar/cgi-bin/koha/opac-detail.pl?biblionumber=87098>, consultada el 18/03/2018 a las 23:35 hs.
- [28] ISO/IEC 25000 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) – Overview
- [29] Ceballos, J.: Caso de Éxito de certificación de producto software según las normas serie ISO/IEC 25000, NEWSLETTER IRAM AÑO 9, N° 72, 2015.

DISEÑO Y CONSTRUCCIÓN DE SISTEMAS DE IOT SEGUROS Y ESCALABLES

Sebastián U. Flores, Mario Berón, Daniel Riesco, Pedro Rangel Henriques
Departamento de Informática - Facultad de Ciencias Físico Matemáticas y Naturales - Universidad
Nacional de San Luis - Ejército de Los Andes 950 - San Luis – Argentina
Universidade do Minho Braga - Portugal
sebastian.ur.flores@gmail.com, {mberon,driesco}@unsl.edu.ar
pedrorangelhenriques@gmail.com

RESUMEN

En la actualidad, tanto los avances en las tecnologías de la información y de la comunicación como la reducción de costos y tamaño de los dispositivos electrónicos, han impulsado el surgimiento de nuevos servicios que conectan el entorno, con redes informáticas. Estos progresos han posibilitado la extracción de información y la realización de cambios en el entorno, a través de diversos tipos de sensores y actuadores. Esta clase de servicios conforman una nueva rama de las tecnologías conocida como IoT, que abarca una amplia cantidad de áreas de investigación y que busca fundamentalmente una mejora en la calidad de vida de las personas, un incremento de la eficiencia de los procesos industriales y un mayor cuidado del medio ambiente y de sus recursos, transformando para ello la forma en que interactúan las personas con el entorno.

En este artículo, se presenta una línea de investigación que aborda el diseño y construcción de *sistemas de IoT Escalables y Seguros*. Un sistema de IoT está conformado por personas y por dispositivos compuestos por sensores y actuadores, en el cual las partes componentes interactúan entre sí, manteniendo un cierto grado de interdependencia para el correcto funcionamiento del sistema. La heterogeneidad antes mencionada hace que la escalabilidad y seguridad de los sistemas de IoT presenten importantes desafíos de investigación.

PALABRAS CLAVE

Sistema de IoT, Dispositivo IoT, Escalabilidad, Seguridad, Privacidad, Arquitectura, Machine to Machine.

CONTEXTO

La presente línea de investigación se enmarca en dos Proyectos de Investigación. El primero: “Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad” – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N.º P-031516. Tal proyecto es la continuación de diferentes proyectos de investigación, a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el Programa de Incentivos. El segundo proyecto: “Fortalecimiento de la Seguridad de los Sistemas de Software mediante el uso de Métodos, Técnicas y Herramientas de Ingeniería Reversa” realizado en conjunto con la Universidade do Minho Braga, Portugal, fue recientemente aprobado por el Ministerio de Ciencia, Tecnología e Innovación Productiva (MinCyT), y su código es PO/16/93.

1. INTRODUCCIÓN

El término IoT es la abreviación de la frase en inglés “Internet of Things” (Internet de las Cosas). Este término corresponde a un dominio de aplicación que integra diferentes campos tecnológicos y sociales. De acuerdo con la IEEE [1], a pesar de que en los últimos años se han realizado grandes avances en este dominio de aplicación, aún no se ha alcanzado un consenso en cuanto a una definición del mismo, que contenga todas sus características y pueda facilitar una mejor comprensión de esta tecnología emergente.

De acuerdo a lo expuesto por la IEEE en [1], el software es un factor muy importante para los sistemas de IoT:

- Los sistemas operativos de IoT están diseñados para ejecutarse en componentes de pequeña escala de la manera más eficiente posible, mientras que, al mismo tiempo, proporcionan funcionalidades básicas para simplificar y respaldar los objetivos y propósitos de las aplicaciones [6,8,15].
- Al contar con una capacidad de procesamiento reducida y con escaso almacenamiento, es muy importante el desarrollo de interfaces de programación de aplicaciones (APIs) que fomenten la reutilización de componentes y una adecuada gestión de los datos que minimice los costos de almacenamiento y de procesamiento [6,7,9,10].
- Los sistemas de IoT, potencialmente pueden crecer y llegar a componerse por millones de dispositivos diferentes, cada uno ubicado en regiones remotas del planeta y con usuarios ubicados también remotamente. En este contexto, la autogestión y auto-optimización de cada dispositivo y / o subsistema individual pueden ser requisitos importantes. En otras palabras, los comportamientos autónomos podrían convertirse en la norma en sistemas de IoT grandes y complejos [2,6,8,10,13].
- Debido a que los sistemas de IoT producen y manejan información reservada, garantizar la seguridad y privacidad de los datos debe ser considerado prioritario desde los comienzos de su planificación e implementación, de forma que ambas estén integradas en cada proceso del sistema [4,10,11,12,13,14].
- Para garantizar la escalabilidad de un sistema de IoT, este debe poseer un diseño arquitectónico adecuado. Por la característica previamente mencionada, se entiende la creación de un sistema flexible que permita interconectar tantos dispositivos IoT como sea necesario. Dicha conexión se realiza sin que importe el medio de conexión físico o el sistema operativo que posea cada uno de ellos, siempre y cuando utilicen las interfaces de software y protocolos de comunicación adecuados [2,5,6,13].

- Finalmente, los sistemas de IoT pueden tener diferentes niveles de impacto en las personas y la sociedad en la que viven, por lo que deben ser concebidos y conducidos dentro de las restricciones y regulaciones de cada país.

Algunos ámbitos de aplicación de sistemas de IoT son los siguientes:

- Sistemas de domótica.
- Extracción y análisis automáticos o semi automáticos de datos en líneas de producción industrial, que permitan optimizar los procesos de negocio [3].
- Dispositivos capaces de ser vestidos por personas o animales, que permitan monitorear el estado físico de los mismos o que les permitan extender sus capacidades físicas (el término más común asociado a esta clase de dispositivos, a nivel mundial, es el de wearables).
- Dispositivos de monitoreo de cultivos y riego automático e inteligente.

Los sistemas de IoT no necesariamente requieren de interacción humana en su funcionamiento. ETSI [2] trata esta clase de interacciones al definir las comunicaciones “*Machine-to-Machine (M2M)*” (en español, comunicaciones Máquina a Máquina). Las comunicaciones M2M son aquellas realizadas directamente entre dispositivos y / o subsistemas de IoT, con interacción humana escasa o nula. Esta clase de comunicación posee varias aplicaciones:

- Automatización de decisiones de carácter determinístico, las cuales se encuentran totalmente estudiadas y no requieren de intervención humana.
- Aprendizaje de patrones de comportamiento humano o patrones en el funcionamiento de procesos propios del sistema. En este caso, la única interacción humana que podría presentarse es la del encargado de supervisar los resultados del aprendizaje realizado por el sistema.
- Autonomización de un sistema aislado, localizado en una ubicación de difícil acceso para seres humanos.
- Edificios inteligentes que se adapten automáticamente a las condiciones del entorno.

En este caso, existiría una constante comunicación entre subsistemas de IoT distribuidos en todo el edificio y sus alrededores, cada uno evaluando las condiciones de su propio entorno y comunicando los cambios relevantes al resto de los subsistemas, para que cada uno se adapte de la mejor forma a las mismas.

Por todo lo mencionado en los párrafos anteriores, se puede notar que es difícil la creación de un marco general de trabajo que permita desarrollar e integrar sistemas de IoT de diversa escala y con ámbitos de aplicación heterogéneos. Este marco debe ser lo suficientemente flexible como para permitir todo tipo de modificación estructural del sistema, de una forma comprensible y segura. También, debe permitir la obtención de múltiples vistas que permitan analizar diferentes propiedades del sistema. Para crear este marco de trabajo, es necesario detectar y comprender los problemas a los que se enfrentan los arquitectos y desarrolladores de sistemas de IoT y que, de acuerdo con Kranz et al. [3], llevan a que un 60% de las iniciativas IoT se estancan en la fase de prueba de concepto, y del 40% restante, sólo el 26% sean consideradas un completo éxito.

Algunos de los problemas antes mencionados son los siguientes:

- Crecimiento organizado del sistema, pudiendo este gestionar a tantos dispositivos y usuarios como sea necesario, sin importar la ubicación geográfica de los mismos y, minimizando los costos de desarrollo y los riesgos asociados al crecimiento del sistema [5,7].
- Establecimiento de jerarquías entre dispositivos, con el fin de que las comunicaciones se realicen de forma controlada y facilitando el posterior procesamiento de la información comunicada.
- Establecimiento de jerarquías de usuarios, asignando a cada uno un determinado nivel de acceso a los dispositivos del sistema de IoT y a la administración del mismo.
- Soporte por software que simplifique la conexión entre dispositivos con múltiples interfaces de hardware y protocolos diferentes [2].

- Soporte para visualizar el flujo de la información en el sistema de IoT, de una forma global, o local de cada subsistema que lo conforme.

Para concluir esta introducción, es importante mencionar que se prevé un amplio crecimiento en la cantidad de objetos conectados a internet a través de sistemas de IoT. Kranz et al. estima que, en el año 2020, se alcanzará una cifra de 50.000 millones de objetos conectados en todo el mundo [3].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En la actualidad, existe una gran variedad de sistemas operativos destinados a dispositivos IoT (e.g. TinyOS, Linux, RIOT, CONTIKI, Windows 10 IoT Core) [1,15]. También, existe una gran cantidad de librerías y aplicaciones implementadas con el fin de brindar soporte a desarrolladores de sistemas de IoT.

De acuerdo con IEEE [1], frente al crecimiento en la cantidad de *cosas conectadas a internet* (i.e. dispositivos IoT), cada una perteneciendo posiblemente a diferentes *dominios de administración* (i.e. diferentes sistemas/subsistemas de IoT), es necesario repensar por completo los enfoques tradicionales en el desarrollo de aplicaciones web, siendo prioritarios elementos como la escalabilidad y la lógica distribuida.

Coincidiendo con IEEE [1], mientras que IoT promete una mejor vida a través de dispositivos conectados y de la información y métricas que ellos generen, también marca el comienzo de una nueva era en cuanto a la privacidad y la seguridad. Según OWASP [4], uno de los nuevos riesgos de seguridad asociados con IoT, son los llamados *ataques de superficie*, que explotan las fallas de seguridad de los sistemas de IoT para vulnerar a empresas, gobiernos, organizaciones y a usuarios particulares.

Frente a lo mencionado en los párrafos anteriores, está claro que los desarrolladores de sistemas de IoT se enfrentan a un problema principal: *la creación de un sistema de IoT escalable y seguro*.

Esta línea de investigación propone:

- Hacer una investigación profunda sobre las vulnerabilidades de seguridad que afectan a los sistemas de IoT.
- Investigar el estado del arte de los patrones/estilos arquitectónicos de sistemas de IoT, y de los protocolos utilizados en su desarrollo.
- Integrar las investigaciones realizadas con el fin de crear un conjunto de herramientas que faciliten al desarrollador la creación de sistemas de IoT escalables y seguros, brindando soporte en la integración de subsistemas que posiblemente estén ubicados en diferentes posiciones geográficas, posean diferentes sistemas operativos y utilicen diferentes protocolos de software y hardware.

3. RESULTADOS OBTENIDOS/ESPERADOS

Con el objetivo de evaluar los conceptos investigados, se realizó un sistema de prueba, compuesto por los siguientes componentes:

- Una aplicación para Windows, destinada a los usuarios del sistema de IoT. Esta aplicación provee un sistema automatizado de búsqueda que muestra en pantalla todos los dispositivos IoT conectados en la red Wi-Fi local. A continuación, ofrece al usuario la posibilidad de diseñar la arquitectura del sistema de IoT, utilizando todos los dispositivos que fueron encontrados en la búsqueda. Durante esta etapa, el usuario puede:
 - Establecer comunicaciones M2M [2] entre los distintos dispositivos IoT detectados.
 - Definir cuales datos deben ser transmitidos periódicamente desde los dispositivos hacia la aplicación del usuario.
 - Determinar cuáles datos deben ser almacenados por los dispositivos para la obtención de métricas.

Una vez finalizado el diseño arquitectónico, la aplicación automáticamente establece conexión con los dispositivos y les transmite la configuración establecida por el usuario desde la aplicación.

- Una placa Arduino conectada a varios sensores, junto con un transmisor/receptor Wi-Fi,

para que esta pudiera conectarse a la red Wi-Fi local. A continuación, se la programó para que sea compatible con los protocolos utilizados dentro del sistema de IoT y se la conectó a la red Wi-Fi local. Por último, se ejecutó la búsqueda automática en la aplicación del usuario y se comprobó que la placa Arduino se conectó exitosamente al sistema de IoT.

- Dado que se contaba con una única placa Arduino y se deseaba probar la configuración del sistema con más de un dispositivo IoT, se creó un script en Python que simula la ejecución de tales dispositivos. Cada vez que se ejecuta el script, se obtiene un nuevo dispositivo IoT conectado a la red Wi-Fi local, preparado para interactuar utilizando los protocolos del sistema. Se ejecutó nuevamente la búsqueda automática en la aplicación del usuario, pudiendo comprobar que todos los dispositivos simulados se conectaban exitosamente al sistema.
- Se definieron protocolos para la conexión y desconexión entre componentes del sistema de IoT (e.g. dispositivos IoT, aplicaciones de usuario), el establecimiento de configuraciones desde la aplicación de usuario hacia los dispositivos IoT, y la transmisión de mensajes básicos (e.g. el valor medido por un sensor, la asignación de una orden a un actuador).

A futuro se poseen los siguientes objetivos:

- Realizar una investigación profunda acerca de las amenazas de seguridad que afectan a los sistemas de IoT.
- Extender el conjunto de protocolos para cubrir todos los procesos posibles de un sistema de IoT, de una forma que garantice la seguridad del sistema y la privacidad de sus datos.
- Crear patrones de diseño arquitectónico que puedan ser usados en diferentes ámbitos de aplicación de sistemas de IoT e indicar, para cada patrón de diseño arquitectónico, aquellos ámbitos en los que sea más viable su implementación.
- Ampliar las capacidades de la aplicación de usuario, permitiendo al mismo obtener diferentes vistas sobre el estado del sistema y

conectarse con dispositivos, no solo en una red local, sino que también a través de internet y de otras interfaces de conexión (e.g. Bluetooth, ZigBee, Serial).

- Permitir el establecimiento de comportamientos más complejos en los dispositivos IoT, durante la etapa de diseño arquitectónico del sistema, evitando la necesidad de tener que re-programar cada uno de los dispositivos manualmente.

4. FORMACION DE RECURSOS HUMANOS

Los progresos obtenidos en esta línea de investigación sirven como base para el desarrollo de tesis de posgrado, ya sea de doctorado o maestrías en Ingeniería de Software y desarrollo de trabajos finales de las carreras Licenciatura en Ciencias de la Computación, Ingeniería en Informática e Ingeniería en Computación de la Universidad Nacional de San Luis, en el marco de los Proyectos de Investigación.

5. BIBLIOGRAFÍA

- [1] IEEE. Towards a Definition of the Internet of Things (IoT). 27 05 2015. https://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Revision1_27MAY15.pdf
- [2] ETSI. ETSI Technical Specification, "Machine-to-Machine Communications (M2M); M2M Service Requirements." Technical Specification. 08 2010. www.etsi.org/deliver/etsi_ts/102600_102699/102689/01.01.01_60/ts_102689v010101p.pdf
- [3] M. Kranz. Internet Of Things. Construye nuevos modelos de negocio. LID Editorial, 2017.
- [4] OWASP, IoT Attack Surface Areas Project, https://www.owasp.org/index.php/OWASP_Internet_of_Things_Project#tab=IoT_Attack_Surface_Areas
- [5] D. Zeng, S. Guo y Z. Cheng. The Web of Things: A Survey (Invited Paper). 2011. <http://www.jocm.us/index.php?m=content&c=index&a=show&catid=50&id=134>
- [6] IETF, The Internet of Things - Concept and Problem Statement, 2010. <https://www.ietf.org/archive/id/draft-lee-iot-problem-statement-05.txt>
- [7] A. Bassi, M. Bauer, M. Fiedler, T. Kramp, R. v. Kranenburg, S. Lange y S. Messner. Enabling Things to Talk: Designing IoT Solutions with the IoT Architectural Reference Model. *SpringerOpen*. 2013.
- [8] CERP-IoT. Visions and Challenges for Realising the Internet of Things. 2010. http://www.internet-of-things-research.eu/pdf/IoT_Clusterbook_March_2010.pdf
- [9] H. Ren, H. Li, Y. Dai, K. Yang y X. Lin. Querying in Internet of Things with Privacy Preserving: Challenges, Solutions and Opportunities. 13 03 2018. <http://ieeexplore.ieee.org/document/8315210/>
- [10] T. Choudhury, A. Gupta, S. Pradhan, P. Kumar y Y. S. Rathore. Privacy and Security of Cloud-Based Internet of Things (IoT). 28 10 2017. <http://ieeexplore.ieee.org/document/8307328/>
- [11] M. El-hajj, M. Chamoun, A. Fadlallah y A. Serhrouchni. Taxonomy of authentication techniques in Internet of Things (IoT). 12 2017. <http://ieeexplore.ieee.org/document/8305419/>
- [12] J. Singh, T. Pasquier, J. Bacon, H. Ko y D. Evers. Twenty Security Considerations for Cloud-Supported Internet of Things. *IEEE Internet of Things Journal*. 2015.
- [13] A. Zanella, N. Bui, A. Castellani, L. Vangelista, M. Zorzi. Internet of Things for Smart Cities. *IEEE Internet of Things Journal*. 02 2014. <http://ieeexplore.ieee.org/document/6740844/>
- [14] S. Pérez, J. L. Hernández-Ramos, S. N. Matheu-García, D. Rotondi, A. F. Skarmeta, L. Straniero y D. Pedone. A lightweight and flexible encryption scheme to protect sensitive data in Smart Building scenarios. 02 2018. <http://ieeexplore.ieee.org/document/8279412/>
- [15] S. Pérez, P.Gaur, M. P. Tahiliani. Operating Systems for IoT Devices: A Critical Survey. *2015 IEEE Region 10 Symposium*. 05 2015. <http://ieeexplore.ieee.org/document/7166231/>

Análisis y Gestión de Riesgo en Proyectos Software

Un nuevo modelo integrando la metodología SEI y Magerit2

Caballero S. D., Kuna H.D.

Facultad de Ciencias Exactas Químicas y Naturales / Universidad
Nacional de Misiones

{sergiocaballero@gmail.com, hdkuna@gmail.com}

Resumen

Hablar de riesgos es hablar de futuro, de probabilidades, de incertidumbre, de avances o retrocesos.

Toda actividad implica un riesgo, y aunque algunos ubican la etimología en la palabra en *risco*, esto no implica de por sí una valoración negativa, todo cambio implica peligros, pero el peor peligro es la inmovilidad.

Un Riesgo es la probabilidad que ocurra una pérdida. Los riesgos técnicos del software son la medida de la probabilidad y severidad de que se produzcan efectos adversos en el desarrollo, adquisición, mantenimiento etc. de sistema.

Todas las áreas en el desarrollo de sistemas son fuentes potenciales de riesgos de software.

Debido a la importancia de estos, en los proyectos software, se realizará una investigación de distintas metodologías para detectar, analizar, eliminar o minimizar los posibles riesgos a los cuales se somete un proyecto software de pequeña a mediana envergadura durante su ciclo de vida (Planificación, análisis, desarrollo, implementación y mantenimiento) y los riesgos generales a los que se somete el TI¹ de estas organizaciones. Basados en metodologías de análisis y gestión de riesgos estándares de la TI se busca lograr un método ágil, flexible, rápido y sencillo de utilizar por organización que no cuentan con los medios para poder afrontar los altos costos de los estudios de análisis y

gestión de riesgos realizados por consultoras privadas o por soporte informático especializado. Y como resultado de la investigación se generará un software prototipo basado en las mejores técnicas de las metodologías investigadas.

Palabras clave: Análisis y Gestión de Riesgos, Auditoría, Ingeniería del Software

Contexto

Trabajo de Investigación para la generación de la Tesis de Maestría en Tecnologías de la Información UNNE-UNAM.

Introducción

El objetivo de esta investigación es generar un método de análisis y gestión de riesgos adaptando los requerimientos y necesidades informáticas y tecnológicas de las organizaciones pyme, se busca crear un método simple, ágil y de económica implementación, utilizando como base de estudio dos de las metodologías más importantes de análisis y gestión de riesgos en TI[1], utilizando la metodología SEI² CRM³ y adaptando técnicas y elementos de la metodología Magerit V3; para facilitar y automatizar la alta carga de trabajo, gestión, control y mantenimiento de proyectos basados en el método creado, en el marco de esta investigación se realizará una herramienta software. El método y la herramienta servirán para analizar y gestionar los posibles riesgos que pueden

¹ Tecnología de la informática

² Software Engineering Institute - Carnegie Mellon University

³ Continuous Risk Management

tener los activos; generar un procedimiento de seguimiento de los planes de acción de los riesgos gestionados, comunicar y registrar los incidentes ocurridos para evaluar posteriormente el nivel de los ajustes a realizar en el AGR⁴, esto ayudará a eliminar o minimizar los incidentes ocurridos, además, resguardar los activos de las organizaciones para que estos no corran riesgos desconocidos o que su impacto sea mínimo y controlado.

Líneas de investigación y desarrollo

El uso de la TI ha crecido considerablemente en los últimos años y las organizaciones cada vez dependen más de ella para garantizar el éxito en el entorno de negocios actual; la habilidad que tenga la organización para implantar las tecnologías modernas que soporten de manera eficiente y controlada a los procesos de negocio críticos, tiene un gran impacto en su grado de competitividad[2].

Los planes estratégicos de negocio actualmente incluyen iniciativas que involucran la optimización de los recursos informáticos para asegurar la consecución de los objetivos de la organización; como consecuencia de lo anterior, los altos ejecutivos están cada vez más alertas sobre la forma en que la tecnología soporta al negocio y dependen cada día más en los Directores de Tecnología de Información para optimizar la organización. [3]

Este incremento ha añadido complejidad a las arquitecturas tecnológicas y a los procesos para su implantación y administración; por consiguiente, se presentan nuevos riesgos que deben ser mitigados de forma

efectiva y eficiente para mantener el cumplimiento de los objetivos de control. Dichos riesgos se encuentran en su mayoría inmersos en los cada vez más complejos sistemas de cómputos, recursos humanos en la etapa de su desarrollo, implementación y mantenimiento, y de TI en general.

Podemos decir que Gestión de Riesgos es una metodología en la cual se utiliza procesos, métodos y herramientas para gestionar los riesgos encontrados en un proyecto software; posee también métodos específicos para identificar riesgos importantes y estrategias para gestionarlos. Además provee de un entorno disciplinado para la toma de decisiones de una manera proactiva basándose constantemente en identificar que puede salir mal; “la Gestión de Riesgos es importante debido a que ayuda a evitar desastres, re-trabajo y sobre-trabajo, pero aún mas importante, porque estimula la generación de situaciones del tipo ganar-ganar”[4].

Las tareas del análisis y gestión de riesgos no son un fin en si mismas, si no que se adaptan en la actividad continua de gestión de la seguridad. El análisis de riesgos permite determinar como es, cuanto vale y cómo de protegidos se encuentran los activos. [5]

Una de la metodología elegida para el desarrollo de la gestión de Riesgos es la propuesta por el SEI 5 la cual cuenta con los siguientes etapas [6]:

- Inventario de activos
- Propósitos y Objetivos del análisis de riesgos
- Equipo de Trabajo
- Taxonomía de Riesgos
- Estimación de la probabilidad
- Estimación del impacto

⁴ Análisis y Gestión de Riesgos

⁵ Software Engineering Institute

- Exposición al riesgo
- Gestión de los Riesgos
 - Plan de Acción
 - Plan de Contingencia

Etapas del Modelo de Gestión de Riesgos SEI-CRM [7]

1° Identificación: Permite anticipar los riesgos antes de que estos ocurran y se transformen en problemas serios afectando adversamente el desarrollo del proyecto. Cabe destacar, que la identificación de riesgos (al igual que cada una de las etapas del modelo) convendría sea realizado de manera disciplinada y consistente, estimulando a los miembros del equipo de proyecto a formular sus inquietudes y facilitando el análisis posterior.

2° Análisis: Pretende transformar una serie de datos que han sido obtenidos en la etapa de identificación, en información que permita llevar adelante una toma de decisiones enfocada en los riesgos más importantes para el proyecto. “El análisis de riesgos es un proceso sistemático de estimación de la probabilidad de ocurrencia y la magnitud de la pérdida o impacto de cada uno de los riesgos identificados mediante el cual se logra reducir la incertidumbre de la medida y del resultado del acontecimiento asociado a un riesgo.” [8]

3° Planificación: Convierte a la información relacionada a cada riesgo, en medidas y acciones efectivas en un tiempo inmediato y futuro. Esta fase incluye el proceso de acciones para cada uno de los riesgos en particular, como así también, otorgar un rango de prioridad a las acciones y a la implementación de un procedimiento de administración integral de riesgos.

4° Seguimiento: Radica en monitorear continuamente el estado de los riesgos y las acciones que fueron adoptadas, con el objetivo de evitar o reducir las pérdidas. Realizar un

seguimiento de los riesgos, conlleva inevitablemente a tener que tomar una serie de medidas vinculadas con la gestión, haciendo posible a los referentes de la administración del proyecto realizar una permanente y precisa revisión de los planes.

5° Control: Proporciona la factibilidad de corregir las desviaciones que puedan causarse a los planes de gestión efectuados.

La comunicación y documentación del proceso son significativas en el modelo, ya que la carencia de garantías en ellas imposibilita la aplicación de cualquier estrategia de administración. La comunicación está y se hace notoria en el modelo en distintos niveles: el primero de ellos, establece la comunicación que corresponde cumplir entre cada uno de las fases del proceso, un segundo nivel esta determinado por la comunicación dentro del equipo de proyecto y el tercer nivel por la que surge entre el proyecto y los diferentes participantes del mismo[9].

La otra metodología elegida para el trabajo es la metodología MAGERIT V3 [10][11]

La metodología Magerit (Metodología de Análisis y gestión de riesgos de los sistemas de información), es una metodología propuesta por el Ministerio de administraciones Públicas del Gobierno Español para los organismos públicos de este país, realizado por el Ministerio de Administraciones públicas, Centro Criptográfico Nacional y la Universidad Politécnica de Madrid y liberado para su ser utilizado en cualquier ámbito.

Se podría definir a los objetivos principales de Magerit V3 como:

- Concientizar a los responsables de sistemas de información (dueños del proceso) de la existencia de riesgos y procesos, y la necesidad de detenerlos a tiempo.

- Ofrecer un método sistemático para analizar tales riesgos.
- Ayudar a descubrir y planificar, las medidas eficaces para conservar los riesgos bajo control.
- Apoyar la preparación de la Organización en procesos de evaluación, auditoría, certificación o acreditación, según corresponda en cada caso.
- Fundamentar sólidamente los argumentos que defenderán la toma de decisiones por parte de los directivos de la organización..

Esta metodología en su versión 3 se compone de tres partes, que se describen a continuación

El Método: Es el documento que describe los pasos y tareas básicas a realizar en un proyecto de análisis y gestión de riesgos, proporciona una serie de aspectos prácticos y además describe la metodología desde un punto de vista de tres ángulos.

El Catalogo de Elementos: Es un manual que especifica claramente los elementos utilizados por la metodología, define y clasifica cada uno de estos, incorporando ejemplos sencillos y aclaratorios con respecto a:

- Tipos de activos
- Dimensiones de valoración.
- Criterios de valoración.
- Amenazas típicas sobre los sistemas de información
- Salvaguardas a considerar para proteger sistemas de información.

Guía de Técnicas: Describe las técnicas utilizadas en la guía metodológica. Técnicas específicas para el análisis de riesgos, análisis mediante tablas, análisis algorítmico, árboles de ataque, técnicas generales, análisis coste-beneficio, diagramas de flujo de datos, diagramas de procesos, técnicas gráficas, planificación de proyectos, sesiones de

trabajo (entrevistas, reuniones y presentaciones) y valoración Delphi⁶.

Tomado las 9 etapas de la metodología SEI basados en la taxonomía a contrario con los XXX etapas que utiliza Magerit V3. Y utilizando las características de Magerit V3 como ser:

- Tipos de activos
- Dimensiones de valoración de los activos
- Criterios de valoración de los activos
- Amenazas típicas sobre los sistemas de información
- Salvaguardas a considerar para proteger sistemas de información

Se crea el nuevo método el cual está conformado por las fortalezas de estas dos metodologías de AGR, la simpleza del SEI y el catálogo del MageritV3, acompañando a este nuevo método con una herramienta software para la gestión del mismo.

Resultados obtenidos/Esperados

El resultado de este proyecto es generar un método para el análisis y gestión de riesgos en el TI y una herramienta software para la gestión del mismo, integrando las mejoras prácticas de la metodología provista por el SEI análisis y gestión de riesgos basados en taxonomías y MAGERIT V3. La misma será orientada a las pequeñas y medianas organizaciones, la cuales poseen un presupuesto acotado para la inversión en la AGR. Por lo cual además del método se generará una aplicación software para asistir a la organización en la gestión de los riesgos en el ciclo de vida del proyecto software. Y se tomará como caso de estudio de la aplicación del método al SMAUNaM⁷.

⁶ Delphi Wideband el método de la valoración es una técnica consenso-basada de la valoración para estimar esfuerzo.

⁷ Servicio Médico Asistencial de la Universidad Nacional de Misiones

Formación de recursos humanos

Tesis de Maestría en Tecnología de la Información UNNE – UNaM.

Trabajo de Investigación "Análisis de Riesgo en Proyectos Software adaptados a la realidad tecnológica y Socio económica de la Pcia. de Misiones " – Secretaría de Investigación y Post Grado – FCEQyN

Referencias

- [1] L. Jaunarena and C. Belletti, "Uso de métricas para la gestión de riesgos," 2002.
- [2] Q. Gerard and M. O. Cinneide, "Experiences with Software Product Line Development in Risk Management Software."
- [3] H. D. Kuna, S. Caballero, S. E. Jaroszczuk, and M. J. Miranda, "Plan de riesgos para la implementación, desarrollo y mantenimiento de componentes de web 2.0 en bibliotecas, caso de estudio en una biblioteca especializada," 2008.
- [4] R. Bertone, P. Thomas, D. Taquias, and S. Pardo, "Herramienta para la Gestión de Riesgos en Proyectos de Software."
- [5] H. Thimm, "A continuous risk estimation approach for corporate environmental compliance management," in *2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC)*, 2015, pp. 83–88.
- [6] J. Hill and D. Victor, "The Product Engineering Class in the Software Safety Risk Taxonomy for Building Safety-Critical Systems," in *19th Australian Conference on Software Engineering (aswec 2008)*, 2008, pp. 617–626.
- [7] R. P. Higuera and Y. Y. Haimes, "Software Risk Management," 1996.
- [8] S. D. Maniasi, "Identificación de Riesgos de Proyectos de Software en Base a Taxonomías," Instituto Tecnológico de Buenos Aires, 2005.
- [9] J. Esteves, J. Pastor, N. Rodriguez, and R. Roy, "Implementing and improving the SEI Risk Management method in a university software project," *IEEE Lat. Am. Trans.*, vol. 3, no. 1, pp. 90–97, Mar. 2005.
- [10] Secretaría General de Administración Digital, "PAe - MAGERIT v.3 : Metodología de Análisis y Gestión de Riesgos de los Sistemas de Información." [Online]. Available: https://administracionelectronica.gob.es/pae_Home/pae_Documentacion/pae_Metodolog/pae_Magerit.html#.Wq0hoLOvG01. [Accessed: 17-Mar-2018].
- [11] D. Gaïti, C. IFIP World Computer Congress (19th : 2006 : Santiago, E. Tovar Caro, and V. Vega Zepeda, *Network control and engineering for QoS, security and mobility, V : IFIP 19th World Computer Congress, TC-6, 5th IFIP International Conference on Network Control and Engineering for QoS, Security and Mobility, August 20-25, 2006, Santiago, Chile*. Springer, 2006.

Innovación en Sistemas Software

ANÁLISIS DE NUEVOS ESTÁNDARES PARA LA WEB MÓVIL

Rocío A. Rodríguez, Pablo M. Vera, M. Roxana Martínez, Fernando A. Parra Beltrán

CAETI - Centro de Altos Estudios en Tecnología Informática
 Universidad Abierta Interamericana (UAI)
 Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

{rocioandrea.rodriguez, pablomartin.vera, roxana.martinez}@uai.edu.ar
 {Adolfo.ParraBeltran}@alumnos.uai.edu.ar

RESUMEN

Dado el auge de los dispositivos móviles y lo frecuentemente que resulta que los usuarios utilicen los mismos para acceder a contenidos y servicios web, resulta necesario implementar nuevas tecnologías estandarizadas por el W3C (Consortio Web a Nivel Internacional), que no son tan populares o bien tienen aún baja implementación en los browsers. En la presente línea de investigación y desarrollo se analizan los nuevos estándares para la web móvil y la posible implementación de los mismos.

Palabras clave: Web Móvil, Dispositivos Móviles, API, Estándar, W3C

CONTEXTO

Esta línea de I+D forma parte de los proyectos radicados en el Laboratorio de Algoritmos y Software del Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). El proyecto cuenta con financiamiento asignado y una duración de 2 años siendo este su segundo año de ejecución.

1. INTRODUCCIÓN

Los dispositivos móviles no sólo son masivamente utilizados, por lo cual resultaría importante considerarlos al momento de pensar en acceso a contenidos o servicios web; sino también que en muchos casos su hardware es muy robusto abriendo la posibilidad de nuevas aplicaciones web que

puedan sacar provecho del mismo. Esto no siempre sucede, en algunos casos porque no están disponibles estándares que permitan utilizar un determinado sensor o componente del dispositivo móvil y en otros casos tan sólo porque quienes realizan las aplicaciones web no han implementado los estándares disponibles, “la APIs de vibración es un estándar consolidado desde inicios del 2015 sin embargo las aplicaciones no usan esta API como por ejemplo para poder brindar un feedback al usuario” [1]. Este es un uso de la API de vibración [2] que podría ser aprovechado desde la web dado que ya está el estándar incorporado en los navegadores.

El W3C (Consortio Web a nivel internacional) "... es un organismo altamente reconocido en la industria por sus estándares en cuanto a generación, construcción y uso de contenido web" [3], tiene grupos de trabajo los cuales se centran en la creación de estándares [4]. Las etapas de estandarización son:

- 1) Editor Draft (Editores)
- 2) Working Draft (Grupo de Trabajo)
- 3) Last Call Working Draft (Comunidad)
- 4) Candidate Recommendation (Implementación)
- 5) Proposed Recommendation (Se envía al W3C para una última revisión de sus miembros)
- 6) W3C Recommendation (Aceptada)

Un largo camino recorren las propuestas de los equipos de trabajo hasta poder

transformarse en una recomendación aceptada por el W3C que es el último paso. Los grupos de trabajo tienen publicados desde sus inicios las propuestas [5], es decir se puede acceder desde los primeros borradores hasta a los estándares ya reconocidos. Existe una gran cantidad de estándares del W3C. A modo de ejemplo en la figura 1 se muestra cómo se presenta un estándar con sus secciones: Título, Estado (W3C Recommendation), Fecha, Versiones del documento (link de la versión actual, link de última versión), Editores...

Indexed Database API 2.0
W3C Recommendation, 30 January 2018

This version:
<https://www.w3.org/TR/2018/REC-indexedDB-2-20180130/>

Latest published version:
<https://www.w3.org/TR/indexedDB-2/>

Editor's Draft:
<https://w3c.github.io/indexedDB/>

Previous Versions:
<https://www.w3.org/TR/2017/PR-indexedDB-2-20171118/>

Test Suite:
<https://github.com/w3c/web-platform-tests/tree/master/indexedDB>

Issue Tracking:
[GitHub](#)

Editors:
[Ali Abbas \(Microsoft Corp.\)](#)
[Joshua Bell \(Google Inc.\)](#)

Errata for this document are recorded as GitHub issues.

The English version of this specification is the only normative version. Non-normative translations may also be available.

Figura 1. Estandar del W3C

Continuando con el ejemplo de este estándar particular presentado en la figura 1 [6] puede observarse la implementación en los distintos navegadores (ver figura 2), observándose en este caso que ha sido prácticamente completamente implementado en todos sus aspectos en una gran cantidad de navegadores [7]. Finalmente, los tiempos de implementación en las aplicaciones dependerán de que los desarrolladores puedan conocer el funcionamiento de las APIs del W3C y como es el comportamiento de las mismas en los distintos equipos, dado que puede variar dicho comportamiento. “Las API web están incrementalmente convirtiéndose en una parte integral de la web o de las aplicaciones móviles” [8]. “Las APIs web proveen un enfoque sistemático y extensible para la interacción entre aplicaciones. Un gran número de aplicaciones móviles hacen uso de las APIs web para integrar servicios” [9].

Por lo cual resulta importante contribuir a documentar el funcionamiento de las API con pruebas en diversos dispositivos, formas de implementación para los distintos browsers e incluso en algunos casos sugerir usos que podría dársele a las APIs dentro de una aplicación web.

Feature	Specification	Maturity	Current implementations
Simple data storage	Web Storage		Shipped
Database query/update	Indexed Database API		Shipped
Encrypted storage	Web Cryptography API		Shipped
	WebCrypto Key Discovery		
File download	download attribute in HTML 5.1		Shipped

Figura 2. Implementaciones en los distintos navegadores

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Los ejes principales del trabajo son:

- Análisis de estándares para la web móvil.
- Diseño de pruebas de nuevas tecnologías, análisis de compatibilidad con distintos navegadores.
- Analizar ventajas de encapsular funcionalidades en APIs de alto nivel.
- Analizar ámbitos de aplicación e impacto de uso.
- Generación de Guías de Uso y Buenas Prácticas para los estándares analizados.

3. RESULTADOS OBTENIDOS/ESPERADOS

Al analizar los estándares web se puede advertir que existen retardos importantes entre el momento en que los mismos se publican como estándares consolidados por el W3C hasta el momento en que las aplicaciones aprovechan dichos estándares. Existen dos motivos por los cuales sucede esto, por una parte desde que el estándar se encuentra vigente hasta que los distintos navegadores lo implementan hay un tiempo de demora, luego una vez implementado por los distintos navegadores los desarrolladores web no aprovechan los mismos por desconocimiento de su existencia, falta de conocimiento en cuanto al uso o manejo de una determinada API e incluso por no haber imaginado un uso para el mismo dentro del contexto de una aplicación. Es por ello que muchos de estos estándares no tienen un grado de implementación alto en las nuevas aplicaciones, impidiendo una innovación que podría ser altamente importante y beneficiar a los usuarios finales. En esta línea de investigación y desarrollo, se analizaron diversos estándares obteniendo como resultado documentación e incluso en algún caso el desarrollo de una nueva API en JavaScript que permite encapsular

funcionalidades y simplificar el uso a los desarrolladores.

Se analizaron dos APIs [7], [8] vinculadas con rendimiento: documentándose, comparándose y ofreciendo escenarios de uso. Así también en cuanto a manejo de interfaz, se analizó un estándar vinculado con Eventos Touch [9] analizando la posible identificación de gestos en la pantalla del navegador para facilitar tareas habituales. En cuanto a acceso al hardware se estuvo trabajando con el sensor de proximidad y se construyó una API que encapsula las cuestiones básicas de acceso al sensor de proximidad, permitiendo identificar gestos en el aire y que los mismos desencadenen una acción particular.

La posibilidad de analizar distintos estándares permite colaborar en grupos de trabajo a alumnos tanto de grado como posgrado. Este proyecto está radicado en el CAETI (Centro de Altos Estudios en Tecnología Informática), con una alta participación de alumnos y docentes de la localización centro (ciudad autónoma de buenos aires), así como también de la localización Castelar (provincia de Buenos Aires), extendiendo la participación al resto de las sedes, incluyendo a la Sede Rosario (provincia de Rosario). Esto permite la actualización continua de alumnos y docentes en la tecnología web móvil por medio de la investigación de nuevos estándares.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo está formado por 12 personas.

- 3 Docentes (2 de Postgrado y 1 de Grado).
- 9 (4 Alumnos de Posgrado, 5 Alumnos de Grado).

En el área de dispositivos móviles se encuentran en realización 2 tesis de maestría en la UAI (Universidad Abierta Interamericana) y 1 de doctorado en la UNLP (Universidad Nacional de La Plata).

5. BIBLIOGRAFIA

- [1] Rodríguez, R. A., Vera, P. M., Martínez, M. R., Parra Beltrán, F. A., & Alcidor, J. (2017). Análisis e implementación de nuevas tecnologías para la web móvil. In XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires), 2017. <http://sedici.unlp.edu.ar/handle/10915/62383>
- [2] W3C, "Vibration API (Second Edition)", 2016 <https://www.w3.org/TR/vibration/>
- [3] Martínez, Francisco O., Gustavo A. Uribe, and Fabián L. Mosquera. "OneWeb: web content adaptation platform based on W3C Mobile Web Initiative guidelines." Ingeniería e Investigación 31.1 (2011): 117-126. <http://www.redalyc.org/html/643/64321170014/>
- [4] W3C, "Roadmap of Web Applications on Mobile", 2018 <https://www.w3.org/Mobile/mobile-web-app-state/>
- [5] W3C, AI Standards and Drafts, 2018 <https://www.w3.org/TR/>
- [6] W3C, Indexed Database API 2.0, 2018. <https://www.w3.org/TR/2018/REC-IndexedDB-2-20180130/>
- [7] W3C, "Data Storage", 2018. <https://www.w3.org/Mobile/roadmap/storage.html>
- [8] Bermbach, David, and Erik Wittern. "Benchmarking web api quality." International Conference on Web Engineering. Springer, Cham, 2016. https://link.springer.com/chapter/10.1007/978-3-319-38791-8_11
- [9] Espinha, Tiago, Andy Zaidman, and Hans-Gerhard Gross. "Web api fragility: How robust is your web api client." arXiv, 2014. <https://arxiv.org/ftp/arxiv/papers/1407/1407.4266.pdf>
- [10] W3C, User Timing, 2013 <https://www.w3.org/TR/2013/REC-user-timing-20131212/>
- [11] W3C, Navigation Timing, 2012 <https://www.w3.org/TR/2012/REC-navigation-timing-20121217/>
- [12] W3C, Touch Events, 2013 <https://www.w3.org/TR/2013/REC-touch-events-20131010/>

IMPLEMENTACIÓN DE SOLUCIONES CON RASPBERRY PI ACCESIBLES DESDE DISPOSITIVOS MÓVILES

Daniel Giulianelli, Pablo Cammaro, Rocío Rodríguez, Pablo Vera, Artemisa Trigueros

GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software)
Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Florencio Varela 1903, San Justo, Provincia de Buenos Aires
{dgiulian, pcammarano, rocio.rodriguez, pvera, atrigueros } @unlam.edu.ar

RESUMEN

La Raspberry Pi (RPi) es un SBC (Single Board Computer) que tiene el tamaño de una tarjeta de crédito, desarrollada en Reino Unido por la fundación Raspberry Pi. El modelo más reciente RPi 3 incorpora wifi, a diferencia de los modelos previos a los que era necesario agregarle por USB una pequeña placa externa. El bajo costo y sobre todo su portabilidad ofrecen un atractivo como para desempeñar una gran cantidad de soluciones.

En este proyecto se diseña y desarrollan soluciones que tienen como servidor a la RPi y cuyos clientes serán dispositivos móviles. Debido a las prestaciones y características de la RPi es posible configurarla como servidor de contenido y a su vez como punto de acceso wifi, brindando una solución completa en un tamaño reducido y sin necesidad de contar con hardware o infraestructura de red adicional. Esta característica permite generar una gran cantidad de soluciones algunas de las cuales se presentan en el presente artículo a modo de ilustrar el trabajo que se está realizando en la presente línea de I+D (Investigación y Desarrollo).

Palabras clave: Raspberry, Servidor de Contenidos, Dispositivos Móviles

CONTEXTO

El presente proyecto de Investigación y Desarrollo corresponde al programa PROINCE y es realizado por el grupo de

investigación GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software) perteneciente al Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza (UNLaM). El grupo es interdisciplinar al estar formado por docentes de distintos departamentos y en él se encuentran formándose en actividades de I+D alumnos de la universidad de las carreras de Ingeniería. Siendo este el segundo año de ejecución del presente proyecto.

1. INTRODUCCIÓN

La Raspberry Pi (RPi) así como otras SBC (Single Board Computer), se han hecho muy populares por su alta portabilidad y bajo costo. A fin de poder ejemplificar lo pequeña que es la RPi, en la figura 1 se muestra la RPi3 con una pantalla táctil integrada donde puede apreciarse que tiene menos de 10 cm de longitud. Por estas características ha sido utilizada para realizar aplicaciones en diversas áreas.

La RPi3 [1] tiene cuatro puertos USB, en cuanto a conectividad cuenta con Ethernet (es decir la posibilidad de conectar un cable de red), WIFI y Bluetooth integrado. Cuenta con un procesador de 64 bits y cuatro núcleos y posee un 1 GB de memoria RAM.



Figura 1. Raspberry Pi 3 con pantalla integrada

Se realizó un mapeo sistemático del estado del arte, en distintas fuentes (entre ellas: ACM, IEEE, ResearchGate, Revistas Indexadas), en base a parámetros precisos. Los trabajos obtenidos, de este mapeo sistemático, se organizaron por áreas temáticas considerándose 3 áreas principales: Educación, IoT (Internet de las cosas) y Redes. No obstante, se seleccionaron otros trabajos de interés representados por la última fila de la tabla 1.

Tabla 1. Porcentaje de trabajos, según se incluyen o no dispositivos móviles, categorizados por áreas

Tema	Porcentaje de trabajos	
	Incluyen dispositivos móviles	No incluyen Dispositivos móviles
IoT	37%	19%
Educación	26%	28%
Redes	16%	32%
Otros	21%	21%

Se tomaron en consideración luego los relacionados con educación y tal como se muestra en la tabla 2, puede observarse que la mayor cantidad de trabajos son originarios de India, seguido por Estados Unidos y Reino Unido; dada la gran inserción a nivel mundial

que tuvo la RPi hay diversos países que aportan trabajos de relevancia que están representado en el último renglón de la tabla 2.

Tabla 2. Porcentaje de trabajos seleccionados por país

Origen	Porcentaje no incluye teléfono móvil
India	40,54%
Estados Unidos	20,27%
Reino Unido	6,76%
Otros	32,43%

Los porcentajes mencionados previamente muestran el gran abanico de acción utilizando la Raspberry Pi. En el presente proyecto se plantea utilizar la RPi como servidor para brindar contenidos o servicios a ser consumidos desde dispositivos móviles.

Las ventajas de realizar soluciones utilizando la Raspdberry Pi como punto de acceso, con clientes móviles, son:

- Portabilidad de la solución
- Los clientes de la red pueden utilizar sus dispositivos móviles sin requerir hardware particular para dicho fin
- La solución a generar podrá ser utilizada en ámbitos donde no se cuenta con conectividad internet ni infraestructura de red wifi.
- Implantar la solución en una nueva RPi implicará tan sólo copiar la imagen de la tarjeta de memoria, lo cual ya permite tener el software instalado con toda la configuración necesaria.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Se investigará principalmente sobre los siguientes temas:

- Dispositivos Móviles
- Redes AdHoc
- Software a Medida
- Diseño de Interfaz táctil

- Pruebas de acceso desde Dispositivos Móviles
- Visualización de datos adquiridos en tiempo real

3. RESULTADOS OBTENIDOS

Durante el primer año del proyecto se desarrolló una aplicación que es utilizada en el ámbito académico, con la finalidad de que los alumnos den respuestas a preguntas de opción múltiple que se realizan sobre los temas de la materia en la clase previa al parcial y de esta forma el docente pueda conocer el porcentaje de alumnos que selecciono cada respuesta posible, evidenciando las dificultades de los alumnos, haciendo que estas preguntas sean los disparadores para el repaso a realizar. Es una solución que permite conocer de forma automática los porcentajes sin tener que contar manualmente cuantos alumnos seleccionarían la respuesta correcta, haciendo que estos levanten la mano en aquella opción que crean que es válida (contabilizando el docente cuantos son en cada caso y anotándolo por ejemplo en el pizarrón), por otra parte, no causa que un alumno elija esa respuesta porque observa que muchos de sus compañeros la han elegido.

Se analizaron distintas aplicaciones gratuitas para la elaboración de formularios seleccionándose Testmaker [2]. Sobre esta herramienta se realizaron algunas modificaciones ya que la misma es de código abierto. Una de las modificaciones realizadas fue la visualización de los formularios, para que estos tuvieran un diseño adaptativo que permita la correcta visualización en distintos dispositivos. “El diseño adaptativo (responsive design) es una buena solución para que un sitio web pueda adaptarse a la visualización en un dispositivo concreto (con una resolución inferior, un tamaño de pantalla más reducido, etc). Permitiendo variar la forma en que se muestra el menú principal, cambiando los mecanismos de navegación, etc” [3]. Existe una gran cantidad de

frameworks adaptativos, se ha elegido a W3CSS, un framework liviano [4] que permitió cambiar el aspecto de la solución final. Si bien este framework es más reducido que otros más populares como Bootstrap [5] o Foundations[6], ha permitido realizar todo lo necesario.

Adicionalmente se desarrolló, utilizando la biblioteca Node.js [7], una aplicación que genera un log que permite visualizar los usuarios conectados (ver figura 2) y también gráficos que serán visualizados por el docente en la RPi o en su celular permitiendo conocer los resultados en tiempo real (expresados en porcentajes), dejando claridad sobre las respuestas seleccionadas por los alumnos. Para los gráficos se utilizó Chart.js [8], una biblioteca de código abierto que permite agregar gráficos adaptativos a una página web.

El testeo inicial de la aplicación fue realizado en el GIDFIS (Grupo de Investigación, Desarrollo, Formación e Innovación en Ingeniería de Software), en donde fue probada con 23 dispositivos (ver figura 3), por una parte utilizando computadoras de escritorio, notebooks, tablets, teléfonos celulares e incluso otras 2 RPi.

Fecha y hora de inicio	Máxima cantidad de usuarios	Duración
8/10/2017-09:24:29	3	10 segundos
8/10/2017-09:39:42	7	15 segundos
8/10/2017-20:40:49	2	6 segundos
8/10/2017-20:41:7	5	4 segundos
9/10/2017-10:41:17	10	30 segundos
10/10/2017-19:19:46	8	44 segundos
10/10/2017-19:42:32	23	81 segundos

Figura 2. Captura de un Log con los datos de los usuarios conectados

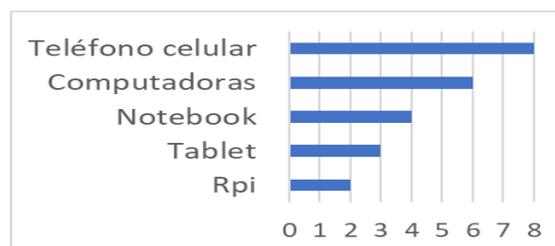


Figura 3. Equipos de pruebas

Actualmente se encuentra en uso esta aplicación. En el segundo año del proyecto, se está trabajando en la implementación de otra aplicación que permitirá acceder a contenidos de forma tal que al insertar un pendrive en la RPi3 se tomarán todos los archivos que estén localizados en una carpeta particular y estos se disponibilizarán mediante una página web para que los alumnos puedan descargarlos. Para esto se está programando un demonio en Linux que chequea el contenido de la carpeta y lo visualiza para que sea consumido por los distintos clientes de la red.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por:

- 10 Docentes pertenecientes a diferentes departamentos conformando un grupo interdisciplinario.
- 3 Graduados (1 Ingeniero en Informática y 2 Ingenieros en Electrónica)
- 2 Alumnos de ingeniería quienes se están formando en actividades de Investigación y Desarrollo los cuales cuentan con becas asignadas.

Vinculado con esta temática se encuentra en realización una tesis de maestría.

5. BIBLIOGRAFIA

- [1] Raspberry Pi 3 Model B.
<https://www.raspberrypi.org/products/raspberry-pi-3-model-b/>
- [2] Testmaker. “Testmaker: Aplicación para crear cuestionarios online”.
<http://jharo.net/dokuwiki/testmaker>
- [3] Rodríguez, R., Vera, P., & Marko, I. (2015). El Gobierno Electrónico y la Implementación de las TIC para Brindar Nuevos Canales de Comunicación. Revista Latinoamericana de Ingeniería de Software, 3(5), 187-196.
- [4] W3Schools. “W3CSS Tutorial”.
<https://www.w3schools.com/w3css/>
- [5] Otto M., Thornton J. “Bootstrap”. 2010 Disponible en: <http://getbootstrap.com/>
- [6] Zurb. “Foundation - The most advanced responsive front-end framework in the world”. 2011. Disponible en: <http://foundation.zurb.com/>
- [7] Node.js Foundation, ” Node.js,”
<https://nodejs.org/es/>
- [8] Charts.js, “Open source HTML5 Charts for your website”
<http://www.chartjs.org/>

INTERFACES TANGIBLES – UNA NUEVA FORMA DE INTERACTUAR CON LOS SMARTPHONES

Pablo M. Vera, Rocío A. Rodríguez, M. Roxana Martínez

CAETI - Centro de Altos Estudios en Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

{pablomartin.vera, rocioandrea.rodriguez, roxana.martinez} @uai.edu.ar

RESUMEN

Las interfases de usuario tangibles (TUI) proponen una forma más natural de manejar un dispositivo. Este trabajo se enfoca en proponer nuevas formas de interacción con los smartphones aprovechando para eso todo el hardware que los mismos tienen disponible. Con hardware no se hace referencia únicamente a la memoria interna, capacidad de almacenaje... sino a un conjunto de sensores y componentes que los mismos tienen. Estos sensores y componentes permitirán enriquecer las aplicaciones e incluso la interfaz de usuario. Si bien es cierto que esto puede favorecer a personas con discapacidades e incluso a quienes no son nativos digitales y no tienen afianzado el manejo de la tecnología, consideramos que el uso de TUI facilita las tareas y el tiempo de ejecución de las mismas lo cual es placentero para todo tipo de usuario que tenga poco tiempo y necesidad de concretar una determinada acción con su teléfono móvil.

Palabras clave: Interfaz, Tangible, Interacción Física, Dispositivos Móviles, Smartphone

CONTEXTO

Esta línea de I+D forma parte de los proyectos radicados en el Laboratorio de Algoritmos y Software del Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). El proyecto cuenta con financiamiento asignado y una duración de 2 años, siendo este el último año.

1. INTRODUCCIÓN

“Los dispositivos móviles forman parte de la vida cotidiana y son cada vez más sofisticados, su poder de cómputo genera posibilidades hasta hace años no pensadas. La creciente demanda de software específico para estos dispositivos ha generado nuevos desafíos para los desarrolladores, ya que este tipo de aplicaciones tiene sus características propias, restricciones y necesidades únicas, lo que difiere del desarrollo de software tradicional” [1]. “Poco a poco nuestra vida se ha ido llenando de pequeños dispositivos pensados para ser usados en movilidad” [2].

El dispositivo móvil por excelencia es el teléfono celular. En este artículo el foco está puesto en los smartphones, los cuales actualmente pueden tener disponible una gran cantidad de componentes y sensores (por ejemplo: Cámara, Micrófono, Sensor de Proximidad, Sensor de Luz, Barómetro, Brújula, Giroscopio, Acelerómetro, GPS, Magnetómetro, Temperatura Ambiente, NFC, Vibración).

Es importante considerar estos componentes y sensores, para enriquecer la aplicación e incluso ofrecer nuevas formas de interacción. “Las interfaces de usuario tangibles (TUI) son interfaces de usuario en las cuales las personas interactúan con información digital a través de ambientes físicos” [3]. Por otra parte “las TUI muestran un potencial para mejorar la manera en que las personas interactúan y aprovechan la información digital” [4].

En diversas ocasiones se pone el foco de atención en la necesidad que las personas aprendan a utilizar las interfaces en vez de considerar si existen otras interfaces que sean más naturales e intuitivas y permitan realizar las mismas acciones con menor tiempo de aprendizaje e incluso de realización de las tareas en cuestión. Si bien se asume que las nuevas generaciones tienen capacidades innatas de utilizar la tecnología, se deja de lado en muchos casos a quienes no son nativos digitales o tienen alguna dificultad física que les impida realizar las tareas con las actuales interfaces. No obstante, si se les propone a los nativos digitales una interfaz natural que permita realizar sus actividades en menor cantidad de tiempo, ellos también se sumarían a la iniciativa. “Los Nativos Digitales se identifican con la interactividad: una respuesta inmediata a todas y cada una de sus acciones” [5].

Algunos autores abordan la necesidad de las TUI para personas con discapacidad física o cognitiva [3], [6], pero así también otros autores las abordan como necesarias para adultos mayores [7], primer infancia [8], [9], [10], [11]... es decir estas interfaces pueden tener un gran potencial para todas las personas, es por ello que algunos autores consideran también su utilización en forma general, independiente de las capacidades físicas, cognitivas, etarias, etc... simplemente por su practicidad y mejora en la concreción de determinadas tareas [12], [13].

Esta línea de I+D (Investigación y Desarrollo) aborda el uso de TUI aprovechando los smartphones y su hardware asociado, considerando que estas interfaces más naturales y cercanas al modelo mental de las personas, podrán facilitar la interacción de aquellos usuarios que tienen capacidades especiales pero también de usuarios que podrían utilizar cualquier tipo de interfaz pero tendrán la alternativa de utilizar el dispositivo de una forma innovadora y más amena.

Para esto se ha elegido desarrollar aplicaciones nativas, por tener estas, acceso

completo al hardware [14]. Si bien las aplicaciones web actualmente tienen acceso a hardware, aún es limitado y para este proyecto en particular es necesario tener acceso a todos los sensores y componentes. Se ha elegido Android como sistema operativo dado que es el sistema operativo más utilizado actualmente. Tomando como referencia las estadísticas proporcionadas por StatCounter [15] se puede apreciar que en Argentina, en base a lo relevado en el mes de enero del 2018, para Teléfonos Celulares el sistema operativo que lidera es Android.

Como puede notarse en la figura 1 en base a los datos estadísticos extraídos [15], la brecha que separa a Android (con el 90%) de los sistemas operativos restantes es muy amplia contando IOS con un 7%, Windows con 2%, mientras que otros sistemas operativos reúnen tan sólo el 1% faltante. Es por ello que se ha decidido en el marco de esta línea de investigación y desarrollo construir aplicaciones nativas para Android.

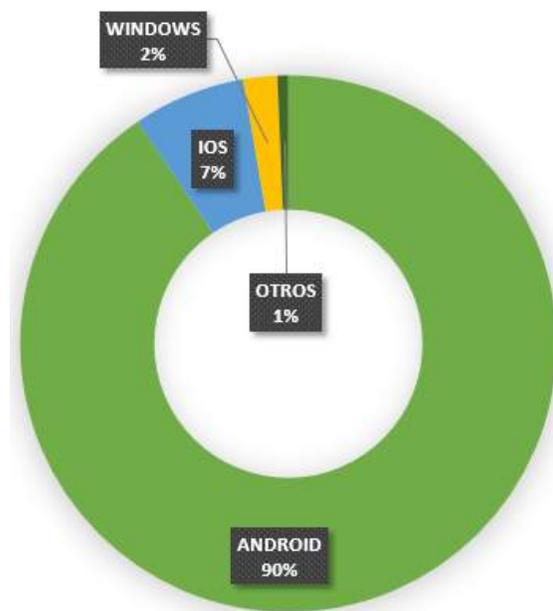


Figura 1. Sistemas Operativos de Teléfonos Celulares – Estadística en Argentina Enero 2018

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Los ejes principales del trabajo son:

- Análisis de los sensores disponibles en los equipos móviles comprendiendo su funcionamiento. De esta forma se podrá determinar las posibilidades y limitaciones al momento de acceder a la información de los sensores mediante una aplicación
- Diseño de interfaces innovadoras que mediante el uso de los sensores permitan manejar distintas aplicaciones.

3. RESULTADOS OBTENIDOS/ESPERADOS

Primeramente, se analizó el funcionamiento de diversos sensores así como los resultados que los mismos arrojaban, entre ellos por ejemplo el sensor de proximidad en donde la sensibilidad del sensor y el tiempo de respuesta varía mucho dependiendo del hardware, con lo cual los resultados son distintos en cada dispositivo. Los resultados obtenidos en algunos casos fueron producto de publicaciones académicas, como es el caso del comportamiento de este sensor y como analizar los valores resultantes [16]. También se ha trabajado con otros componentes como NFC [17], logrando que al apoyar un smartphone a una superficie, incluso sin tener abierta la aplicación esto dispare eventos como una llamada telefónica. Esta aplicación tiene usos particulares como por ejemplo: para personas no videntes que pueden tener un tablero con los nombres en braille de sus contactos haciendo que al apoyar el teléfono sobre la superficie se inicie en forma automática la llamada con el altavoz activado. También puede aplicarse a adultos mayores para que rápidamente puedan comunicarse con sus contactos sin necesidad de manipular el teléfono ganando también velocidad y practicidad en caso de una emergencia.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo está formado por 7 personas.

- 3 Docentes (2 de Postgrado y 1 de Grado).
- 4 (2 Alumnos de Posgrado, 2 Alumnos de Grado).

En el área de dispositivos móviles se encuentran en realización 2 tesis de maestría en la UAI (Universidad Abierta Interamericana) y una de doctorado en la UNLP (Universidad Nacional de La Plata).

5. BIBLIOGRAFIA

- [1] Delía, L., Galdamez, N., Thomas, P., Pesado, P. M. (2013). Un análisis experimental de tipo de aplicaciones para dispositivos móviles. In Congreso Argentino de Ciencias de la Computación (CACIC)(Vol. 18).
- [2] Vázquez, Natalia Arroyo. Información en el móvil. Vol. 4. Editorial UOC, 2011.
- [3] Haro, Barbara Paola Muro, PEDRO CÉSAR SANTANA MANCILLA, and Miguél Ángel García Ruíz. "Uso de interfaces tangibles en la enseñanza de lectura a niños con síndrome de Down." *El Hombre y la Máquina* 39 (2012): 19-25.
<http://www.redalyc.org/pdf/478/47824590004.pdf>
- [4] Shaer, Orit, and Eva Hornecker. "Tangible user interfaces: past, present, and future directions." *Foundations and Trends in Human-Computer Interaction* 3.1-2 (2010): 1-137.
- [5] Marc Prensky "Nativos e Inmigrantes Digitales", 2001.
[https://www.marcprensky.com/writing/Pr-en-sky-NATIVOS%20E%20INMIGRANTES%20DIGITALES%20\(SEK\).pdf](https://www.marcprensky.com/writing/Pr-en-sky-NATIVOS%20E%20INMIGRANTES%20DIGITALES%20(SEK).pdf)

- [6] Ávila-Soto, Mauro, Elba Valderrama-Bahamóndez, and Albrecht Schmidt. "TanMath: A Tangible Math Application to support children with visual impairment to learn basic Arithmetic." Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments. ACM, 2017.
- [7] Galiev, Ruzalin, Dominik Rupprecht, and Birgit Bomsdorf. "Towards Tangible and Distributed UI for Cognitively Impaired People." International Conference on Universal Access in Human-Computer Interaction. Springer, Cham, 2017.
- [8] González González, Carina Soledad. "Revisión de la literatura sobre interfaces naturales para el aprendizaje en la etapa infantil." (2017).
- [9] Xie, Lesley, Alissa N. Antle, and Nima Motamedi. "Are tangibles more fun?: comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces." Proceedings of the 2nd international conference on Tangible and embedded interaction. ACM, 2008.
- [10] Devi, Suraksha, and Suman Deb. "Augmenting Non-verbal Communication Using a Tangible User Interface." Smart Computing and Informatics. Springer, Singapore, 2018. 613-620.
- [11] Bouabid, Amira, Sophie Lepreux, and Christophe Kolski. "Design and evaluation of distributed user interfaces between tangible tabletops." Universal Access in the Information Society (2017): 1-19.
- [12] De Raffaele, Clifford, Serengul Smith, and Orhan Gemikonakli. "Explaining multi-threaded task scheduling using tangible user interfaces in higher educational contexts." Global Engineering Education Conference (EDUCON), 2017 IEEE. IEEE, 2017.
- [13] Anastasiou, Dimitra, and Eric Ras. "A Questionnaire-based Case Study on Feedback by a Tangible Interface." Proceedings of the 2017 ACM Workshop on Intelligent Interfaces for Ubiquitous and Smart Learning. ACM, 2017.
- [14] Rodríguez, Rocío Andrea, et al. "Analysis of current and future web standards for reducing the gap between native and web applications." XX Congreso Argentino de Ciencias de la Computación (Buenos Aires, 2014). 2014.
- [15] StatCounter, GlobalStat. "Mobile Operating System Market Share Argentina". 2018. <http://gs.statcounter.com/os-market-share/mobile/argentina/#monthly-201801-201801-bar>
- [16] Vera, Pablo Martin, and Rocio Andrea Rodriguez. "Creating and Using Proximity Events on Mobile Websites." IEEE Latin America Transactions 14.11 (2016): 4579-4584.
- [17] Rodríguez, Rocío A., et al. "Context Aware Applications on Mobile Environments-Engaged by the use of NFC." International Conference on Multimedia, Scientific Information and Visualization for Information Systems and Metrics. 2013.

Innovación en la gestión de la salud asistencial y laboratorios de investigación en salud

Irrazábal, Emanuel¹; Petrella, Bettina^{2,3}, Cortés, María Alicia^{4,5}

1: Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste

2: Departamento de Informática, Facultad de Ciencias Exactas Químicas y Naturales. Universidad Nacional de Misiones.

3: Jefatura Área de Sistemas. Ministerio de Salud Pública. Misiones

4: Consejo Nacional de Investigaciones Científicas y Técnicas

5: Facultad de Medicina. Universidad Nacional del Nordeste

eirrazabal@exa.unne.edu.ar, bepetrella@hotmail.com, mariaalicia.cortes@gmail.com

Resumen

Esta línea de investigación aborda temas de innovación en los sistemas de gestión de salud. Y hace especial foco en dos temáticas. La primera son los sistemas de gestión para la salud asistencial, especialmente en los centros y las redes de atención primaria. La segunda es la gestión de laboratorios de investigación, tanto en el ámbito público como en el ámbito privado; especialmente en el desarrollo de aplicaciones para la gestión de los protocolos de trabajo.

En ambos casos se espera la aplicación, adaptación, validación y posterior transferencia de las tecnologías actuales al ámbito regional. En particular se está trabajando con el Ministerio de Salud de la provincia de Misiones para el desarrollo de un modelo de gestión del territorio en su red de Centros de Atención Primaria. Asimismo, se está trabajando con la Facultad de Medicina de la Universidad Nacional del Nordeste para la elaboración e instrumentación de un sistema de gestión de biobancos con fines de investigación.

Palabras clave: Innovación, Sistema de Gestión, Salud, CAPS, Biobanco.

Contexto

La línea de Investigación y Desarrollo presentada en este trabajo corresponde al proyecto PI-F17-2017 “Análisis e implementación de tecnologías emergentes en sistemas computacionales de aplicación regional.”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2018-2021.

Asimismo, parte de la línea de investigación es realizada en el marco de la tesis de la maestranda Bettina Petrella perteneciente a la Maestría de Tecnologías de la Información Rs. 764/14 CS UNNE.

Introducción

Gestión de la Salud Asistencial

A nivel mundial, se ha comprobado la importancia de la utilización de la informática en la gestión de las redes de salud, en especial para identificar el área de cobertura de los centros asistenciales [1]. Esto tiene un impacto directo en el territorio, obteniendo datos, controlando y realizando acciones de prevención en la población de referencia [2]; según las características de la zona evaluada [3][4][5][6][7][8][9] en relación con los

diferentes diagnósticos epidemiológicos [8] y grupos etarios [9].

En el nordeste argentino, especialmente en la provincia de Misiones la red pública de Centros de Atención Primaria de la Salud (CAPS), está compuesta por 380 centros que brindan prestaciones de atención primaria, registrándose en promedio unas 2.100.000 de consultas anuales[10]. Pero no es posible catalogar estas consultas de acuerdo a la población de referencia. Esto hace imposible también calcular los diferentes indicadores que aseguren una mayor eficiencia en las políticas de salud pública o la realización de atenciones programadas y preventivas a la población[11][12][13].

Actualmente, la gran mayoría de los centros de atención primaria de la red pública de la provincia de Misiones han definido su área de cobertura, y en el mismo centro puede encontrarse la cartografía en papel. Sin embargo, dicha definición no necesariamente ha seguido procedimientos homogéneos entre centros, y su actualización depende de la decisión del personal de salud local. Por ejemplo, ante la aparición de un nuevo barrio, no existe un procedimiento específico para asignar esta nueva población a un centro determinado. A su vez, desde el nivel central, no se cuenta con información sobre el estado actual de cobertura de cada CAPS.

Gestión de Laboratorios de Investigación

Desde otro enfoque y orientado hacia el desarrollo de investigación en salud uno de los elementos de mayor valor agregado es el desarrollo de biobancos. Los biobancos con fines de investigación biomédica son establecimientos públicos o privados, sin ánimo de lucro, que acogen una o varias colecciones de muestras biológicas de origen humano con fines de investigación biomédica, organizadas como una unidad técnica con criterios de calidad, orden y destino.

Los biobancos, por lo tanto, son indispensables en el avance del reconocimiento de marcadores de enfermedad y progresión, así como también en el descubrimiento de nuevos fármacos[14]. Y dentro de las bases que garantizan el éxito de un biobanco se incluye la implantación de un sistema de calidad [14][15][16]. En este contexto, la aplicación de normas de calidad genéricas permite lograr mejoras continuadas en el tiempo. Entre las normas de calidad genéricas más utilizadas actualmente, se encuentra la norma internacional ISO 9001 que se aplica a los sistemas de gestión de calidad (SGC)[17].

Esta norma se utiliza para implementar y mejorar la eficacia de un SGC, aumentando la satisfacción del cliente mediante el cumplimiento de sus requisitos y es aplicada por más de un millón de organizaciones en 178 países [18]

En cuanto a los laboratorios de investigación, las guías de buenas prácticas para los biobancos publicadas por la Organización para la Cooperación y Desarrollo Económico [19], el Instituto Nacional del Cáncer [20] y la Sociedad Internacional para el Almacenamiento Biológico y Medioambiental [21] aconsejan normativas basadas en ISO 9001 como la ISO 17025 (control y calibración de laboratorios) y la ISO Guide 34 (producción de material). Por otra parte, Betsu et al., analizando los estándares propuestos, recomiendan el uso de la ISO 9001 como norma general aplicable[22]

Finalmente, en los biobancos que han decidido seguir el modelo presentado por la ISO 9001 el foco de la certificación está en la gestión de los procesos que realizan [24][25].

Líneas de investigación y desarrollo

En la línea de Gestión de la Salud Asistencial se propone:

- Construir una metodología de trabajo con guías y procedimientos para la gestión de los CAPS con indicadores poblacionales y la gestión territorial de la red.
- Desarrollar el marco tecnológico para instrumentar la metodología de trabajo.
- Realizar la integración y el ensayo de la metodología, en el Ministerio de Salud de la Provincia, con el marco tecnológico desarrollado.

En la línea de Gestión de Laboratorios de Investigación se propone:

- Desarrollar un modelo de gestión de biobancos de acuerdo con la normativa ISO 9001 y los reglamentos facultativos argentinos / españoles.
- Construir una aplicación para la gestión de redes de biobancos de muestras para investigación en cáncer.

Finalmente, existe una relación directa entre la gestión de la red de CAPS y la gestión de biobancos en red. Un sistema para la gestión de CAPS permite mantener no solamente la información geográfica sino también la información de los diferentes tratamientos de cada paciente. Y este es un conocimiento imprescindible al momento de realizar experimentos con los conjuntos de muestras almacenadas en los biobancos.

Resultados obtenidos

El grupo de investigación es de reciente formación, por lo cual los resultados son preliminares y, en parte, se enumeran antecedentes llevados adelante en el marco de otros grupos de trabajo. A continuación se indican:

En la línea de Gestión de Laboratorios de Investigación:

- Se llevó adelante el desarrollo de los procedimientos para la puesta en marcha de un sistema de gestión de calidad en un nodo de biobanco de

muestras renales de la red europea de biobancos RedInRen [27].

- Se desarrollaron un conjunto de procedimientos compatibles con la norma ISO 9001 para la cesión de muestras del biobanco [28] y se analizó su impacto posterior [29].

Formación de recursos humanos

En el Grupo de Investigación en Innovación en Software y Sistemas Computacionales (GISSC) están involucrados 4 docentes investigadores, 1 becario de investigación de pregrado, 1 tesista de doctorado y 3 tesistas de maestría. Cinco alumnos de la carrera están realizando sus proyectos finales vinculado a estos temas.

Referencias

- [1] M. L. Ramírez, “¿Dónde localizar Hospitales Públicos? Las Nuevas Tecnologías –SIG– como herramientas de apoyo a la planificación territorial. Un caso de estudio aplicado a la Provincia del Chaco -Argentina,” *Número Ser. Geográfica*, vol. 10, pp. 1136–5277, 2001.
- [2] G. D. Buzai, “Tipología de áreas de influencia de los Centros de Atención Primaria de Salud en la ciudad de Luján, Provincia de Buenos Aires, Argentina,” *Huellas*, vol. 20, pp. 35–56, 2016.
- [3] D. A. Lovett, A. J. Poots, J. T. C. Clements, S. A. Green, E. Samarasundera, and D. Bell, “Spatial and Spatio-temporal Epidemiology Using geographical information systems and cartograms as a health service quality improvement tool,” *Spat. Spatiotemporal. Epidemiol.*, vol. 10, pp. 67–74, 2014.
- [4] P. A. Chikumba, “Exploring Integrative Approach of GIS Implementation: The Case of GIS in Health Management in Malawi,” *IIMC International Information Management Corporation, on Conf. 2017 IST-Africa Week*

- Conference , IST-Africa, May 30 2017-Jun 2 2017, pp. 1–9.
- [5] P. Hathi, S. Haque, and L. Pant, “Place and Child Health: The Interaction of Population Density and Sanitation in Developing Countries,” *Demography*, Vol. 54, pp. 337–360, 2017.
- [6] P. T. Makanga *et al.*, “Seasonal variation in geographical access to maternal health services in regions of southern Mozambique,” *Int. J. Health Geogr.*, vol 16 nro 1, pp.1–16, 2017.
- [7] E. Med and O. Investigation, “Assessment of Under-5 Mortality Rates in İstanbul Using the Geographic Information System,” vol. 38, nro. 1, pp. 6–11, 2016.
- [8] J. P. Silva, “Mapping Unhealthy Behavior Among Economically Active Men Using GIS in Suburban and Rural Areas of Sri Lanka”, *Asia Pacific Journal of Public Health*, vol. 28, no 1, pp. 10–16, 2016.
- [9] A. Rautio, J. Rusanen, A. Taanila, T. Lankila, and M. Koiranen, “Is geographical distance a barrier in the use of public primary health services among rural and urban young adults? Experience from Northern Finland”, *Public Health*, vol. 131, pp 82-91, 2016
Ministerio de Salud de la Nación, “Sistema Integrado de Información Sanitaria Argentino”, Julio 2009. [Online]. Disponible: <https://sisa.msal.gov.ar/sisa/#sisa>.
- [10] Dirección de Estadísticas e Información de Salud, “Consultas Ambulatorias En Establecimientos Oficiales-Año 2013”, 2016.
- [11] [E. L. Macquillan, A. B. C. K. M. Baker, and R. P. Y. O. Back, “Using GIS Mapping to Target Public Health Interventions: Examining Birth Outcomes Across GIS Techniques,” *J. Community Health*, vol. 42, no. 4, pp. 633–638, 2017.
- [12] P. R. Ward and A. Regan, “Applying Precision Public Health to Prevent Preterm Birth”, *Front. Public Health*, vol. 5, no. April, pp. 1–16, 2017.
- [13] L. N. Carroll, A. P. Au, L. Todd, T. Fu, I. S. Painter, and N. F. Abernethy, “Visualization and analytics tools for infectious disease epidemiology: A systematic review,” *J. Biomed. Inform.*, vol. 51, pp. 287–298, 2014.
- [14] Carter A, Betsou F. Quality Assurance in Cancer Biobanking. *Biopreserv Biobank* 2011;9(2):157-63.
- [15] García-Merino I, de las Cuevas N, Jiménez J, Gallego J, Gómez C, Prieto C, et al. The Spanish HIV BioBank: a model of cooperative HIV research. *Retrovirology* 2009;6:27.
- [16] Liaño F, Torres AM. Biobancos: una nueva herramienta para la investigación clínica. *Nefrología* 2009;29(3):193-5.
- [17] ISO9001. Quality Management Systems Requirements. Geneva: International Organization for Standardization; 2015.4
- [18] ISO. The ISO Survey of certifications 20145. Available at: <http://www.iso.org/iso/survey2009.pdf> 2010.
- [19] OECD. OECD Best Practice Guideline for Biological Resource Centres (BRCs). Paris; 2007. p. 1-115.
- [20] NCI. National Cancer Institute. First-Generation guidelines for NCI supported Biorepositories. *Federal Register* 2006;71(82):25184-203.
- [21] ISBER. 2008 Best Practices for Repositories Collection, Storage, Retrieval and Distribution of Biological Materials for Research International. *Cell Preserv Technol* 2008;6(1):1-56.
- [22] Betsou F, Luzergues A, Carter A, Geary P, Riegman P, Clark B, et al. Towards norms for accreditation of biobanks for human health and medical research: compilation of existing guidelines into an ISO certification/accreditation norm-

- compatible format. The Quality Assurance Journal 2007;11(3-4):221-94.
- [23] Chaigneau C, Cabioch T, Beaumont K, Betsou F. Serum biobank certification and the establishment of quality controls for biological fluids: examples of serum biomarker stability after temperature variation. Clin Chem Lab Med 2007;45(10):1390-5.
- [24] Nederhand RJ, Droog S, Kluft C, Simoons ML, De Maat MPM; investigators of the EUROPA trial. Logistics and quality control for DNA sampling in large multicenter studies. J Thromb Haemost 2003;1(5):987-91.
- [25] Lacalamita R, Schirone M, Paradiso A. ISO 9001:2000 applied to a research oncology laboratory: which problems? The experience of National Cancer Institute-Bari. Ann Oncol 2008;19(6):1207-8.
- [26] I. García Sánchez, M. Ángel, G. García, and G. Izquierdo (2010). Biobancos: una herramienta necesaria para el futuro de la investigación biomédica. Revista Española de Esclerosis Múltiple. Nº 14.
- [27] Calleros, L., Cortés, M. A., Luengo, A., Mora, I., Guijarro, B., Martín, P., & Rodríguez-Puyol, M. (2012). Puesta en marcha de una plataforma de proceso, almacenamiento y gestión de muestras clínicas: organización y desarrollo del Biobanco de REDinREN. Nefrología (Madrid), 32(1), 28-34.
- [28] Calleros L, Cortés MA, Luengo A, Mora I, Cortés, M. A., Irrazábal, E., García-Jerez, A., Bohórquez-Magro, L., Luengo, A., Ortiz-Arduán, A., ... & Rodríguez-Puyol, M. (2014). Impacto de la implementación de la norma ISO 9001: 2008 en el proceso de cesión de muestras del biobanco Red de Investigación Renal española. Nefrología (Madrid), 34(5), 552-560.
- [29] A. García Jerez, A. Luengo Rodríguez, M. Cortes, E. Irrazabal, L. Calleros Basilio, M. Rodríguez-Puyol. La implantación de la norma ISO 9001:2008 en el biobanco de RedinRen (red de investigación renal): controles de calidad y mejora en el proceso de cesión de las muestras de ADN. Congreso: XLIV Congreso Nacional de la Sociedad Española de Nefrología. Resúmenes XLIV Congreso Nacional de la SEN. ISSN: 1131-4168. Lugar de celebración: Barcelona (España) . 7 de Octubre de 2014

Servicios PaaS de cloud computing e informática contextualizada aplicada a una app híbrida, turística para la ciudad de Puerto Iguazú Misiones

Cristian A. Kornuta

Departamento de Informática, Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones. Félix de Azara 1552. (3300) Posadas, Misiones. Argentina.

{cristian.kornuta24}@gmail.com

Resumen

La plataforma tecnológica de Cloud Computing, es la innovación tecnológica por excelencia de la década, la presente tecnología ha permitido y permite la creación servicios por parte de los gigantes de Internet como ser Mapas en líneas, Base de datos en tiempo real o la creación de sitios como plataformas sociales o plataforma de videos, plataformas que nos permiten transmitir y comunicarnos libremente. Lo anterior permitió que en la actualidad sea tan importante un bloguero en línea como un corresponsal de una cadena de televisión.

Hoy en día surgen innovaciones de nuevos modelos de negocios que se basan en esta tecnología, modelos que hace unos años atrás no se nos hubiese imaginado implementarlo. La tecnología de Cloud Computing conjuntamente con el uso masivo de dispositivos móviles y el concepto de computación contextualizada han propiciado un nuevo camino para todo lo referente a la asistencia en línea de los usuarios. Dentro del sin fin de usuarios de estas tecnologías encontramos a los turistas. Este nuevo turista llamado 2.0 o digital, el cual se convirtió en hiperconectado y multicanal, acostumbrado a usar sus dispositivos móviles se vio beneficiado con la aparición de app móvil que lo hacen sentir como en su casa en el destino donde elige visitar.

Palabras Clave: Cloud Computing – Informática sensible al contexto – Desarrollo móvil.

Contexto

El presente trabajo expone el estudio de tres temas como ser Cloud Computing, Informática contextualizada y el desarrollo de aplicaciones móviles tratados en el curso de Sistemas Distribuidos perteneciente a la Maestría en Tecnología de la Información entre la Universidad Nacional de Misiones y la Universidad Nacional del Nordeste a partir de la revisión bibliografía y material del curso utilizado y referencias externas. Estos temas fueron seleccionados y se desarrollan en el marco del actual proyecto que nos encontramos desarrollando en la ciudad de Puerto Iguazú, Misiones “*Puerto Iguazú móvil*” conjuntamente en la actualidad el presente proyecto se encuentra enmarcado en un proyecto de tesis para la obtención del título en Maestría en Tecnología de la Información por la Universidad Nacional de Misiones y la Universidad Nacional del Nordeste.

Introducción

La plataforma tecnológica de Cloud Computing, es la innovación tecnológica por excelencia de la década, esta tecnología ha permitido y permite la creación servicios por parte de los gigantes de Internet en su mayoría. Servicios que antes solo unos pocos podrían imaginar a un futuro muy lejano. Podemos decir que nos encontramos frente a un nuevo paradigma tecnológico que se impondrá de a poco como lo hizo el internet y la web en su momento afirman. Tanto es el impacto que ha producido esta innovación tecnológica que el concepto de Cloud Computing o “la nube” se

encuentra en boca de los medios de comunicación y de la sociedad en común. La mayoría de las empresas de una u otra forma ya se encuentran utilizando estos servicios y migrando a este nuevo modelo de negocio.

Además, son cada vez más las empresas que en mayor o menor medida ofrecen los mismos servicios que ofrecen los gigantes de internet.

Este nuevo paradigma tecnológica que nos encontramos viviendo vino propiciado por el uso masivo de dispositivos móviles; Dispositivos que permitieron y permiten que el usuario común que antes se encontraba frente a una computadora “de forma fija” o teniendo el problema de tener que desplazar sus documentos de un lugar a otro para poder trabajar en ellos desde otra computadora. A pasar a ser un usuario que la mayoría de las veces de manera gratuita puede desplazarse y tiene acceso a todos los recursos que posee sean estos archivos en su conjunto y de las aplicaciones con las cuales trabaja; con una simple conexión a internet.

No solo esto ha permitido esta tecnología, sino que los prestadores de estos servicios pueden realizar un seguimiento realmente profundo del usuario final con respecto a sus hábitos, consumos y tareas diarias o en que ocupa su tiempo. Lo que llevó a una nueva forma de ofrecer productos y servicios, como ser el caso de Google Maps, o Google Calendar, con el hecho de registrar un evento en algún lugar, como ser la casa por ejemplo y el evento de sacar la basura, él te lo recordará cuando te encuentres en ese lugar porque sabe dónde te encuentras y se lo dijiste.

Todos los días aumenta el número de usuarios en Internet, usuarios que de apoco van formando parte de uno u otro grupo de usuarios que utilizan algún servicios de Internet. Entre estos usuarios beneficiados se encuentran los turistas, los cuales actualmente son llamados turistas 2.0 o llamado digital, el cual se convirtió gracias a estas tecnologías en hiperconectado y multicanal, acostumbrado a

usar sus dispositivos móviles se vio beneficiado con la aparición de app móvil que lo hacen sentir como en su casa en el destino donde elige visitar el uso de estas APP móviles no sólo se da durante el periodo de visita del turista sino antes, durante y después.

Es por eso que este nuevo paradigma de la nube ha impulsado al sector turístico a desarrollar aplicaciones referentes al transporte, gastronomía y cultura por nombrar algunas.

Una de las características que poseen estas aplicaciones, y que será desarrollado en este trabajo, es referente a la informática contextualizada cuyas aplicaciones en el sector es muy amplio.

Otro gran sector beneficiado por los servicios de la nube, y que en mayor o menor medida, son los impulsores de estas nuevas tendencias son los desarrolladores en este sector podemos encontrar Servicios como ser Firebase que es un sistema de Base de datos en tiempo real y con todo el soporte Backend para el desarrollo de aplicaciones.

Línea de investigación y desarrollo

El presente trabajo orienta su línea de investigación en dos grupos; el primer grupo referente al Geo referenciamiento, Geo posicionamiento y al Geo marketing y una segundo grupo a la informática contextualizada como ser la identificación de usuario y el ofrecimiento de servicios o productos orientados a personas específicas.

Objetivos y resultados

Puerto Iguazú es una ciudad 100% turística visitada anualmente por miles de turistas locales, regionales y del resto del mundo, si bien existe mucha promoción a través de los medios de comunicación, folletería e internet; aún no se ha solucionado el problema de la falta de información disponible organizada, concentrada y principalmente actualizada en un solo lugar, que permita al turista mejorar la experiencia en el destino y que el mismo incremente su promedio de estadía en la ciudad y así consuma más; hasta el momento dicho

promedio es de dos días y medio (según IPEC), Iguazú no cuenta con una app para dispositivos móviles que concentre toda la información, basada en la geo localización que permite detectar dónde está el turista y mostrarle todos los puntos turísticos, hoteles, restaurantes, empresas de turismo, entretenimiento nocturno, estaciones de servicio, cajeros automáticos, servicios de salud y todo lo que pueda ser de interés para un turista. La geo localización juntamente con las redes sociales es una nueva forma de hacer marketing, tienen un gran potencial en el turismo ya que el viajero está constantemente conectado y le gusta compartir sus experiencias en las redes sociales (fotos, el lugar donde se encuentran). Existen aplicaciones como:

- **Foursquare:** servicio basado en geo localización web aplicada a las redes sociales.
- **Tripadvisor:** es una aplicación móvil que permite conocer puntos turísticos, restaurantes y hoteles en un determinado lugar a través de la geo localización y el geo marketing.

Ninguna de las aplicaciones mencionadas anteriormente soluciona el 100% del problema expuesto, y la mayor desventaja que presentan es que no proporcionan información actualizada y oportuna para el turista.

“Iguazú en tu móvil” pretende por un lado ser una aplicación que permita al turista vivir una experiencia satisfactoria en su estadía en la ciudad y hacerlo sentir como un local y por otro lado tiene como objetivo principal contribuir al desarrollo económico de la ciudad, a través, del trabajo en conjunto de los actores principales: la municipalidad de Iguazú, la Asociación Civil Atractivos turísticos de Iguazú, los diferentes alojamientos, restaurantes, comerciantes, transporte y otros. Ellos proveerán periódicamente y constantemente la información para la aplicación (promociones, espectáculos etc.). Los actores antes

mencionados serán los encargados de promocionar la aplicación a través de publicidad utilizando cartelería, folletería y las TICs.

Con respecto al desarrollo de la aplicación y mantenimiento estará a cargo del equipo responsable del presente proyecto.

Objetivo general

Lograr a través del desarrollo de una APP con geo localización (en varios idiomas) concentrar, organizar y mantener actualizada toda la información turística y de servicios disponible de la ciudad de Puerto Iguazú para aumentar la oferta turística y así incrementar el promedio de estadía y el consumo del turista en el lugar, beneficiando y contribuyendo de esta manera el desarrollo económico de los comerciantes, empresarios locales y todos los actores involucrados en el proyecto.

Objetivos Específicos

- Promocionar los otros puntos turísticos de Iguazú.
- Hacer sentir al turista como un local.
- Proporcionar al turista no solo una guía turística sino también una guía de servicios e información actualizada sobre promociones, espectáculos y otros.
- Ofrecer recomendaciones personalizadas sobre restaurantes, alojamientos y actividades a realizar.
- Contribuir al desarrollo económico de las empresas y comercios locales relacionados con el turismo.
- Brindar una solución al problema urbanístico de la ciudad. (Posee muchas calles diagonales).
- Comprometer a la Municipalidad de Puerto Iguazú, ACATI, Empresas hoteleras, turísticas, restaurantes, comerciantes, empresas de servicios, expertos en idiomas y otros a trabajar en conjunto.
- A continuación se presentara las pantallas que fueron desarrolladas en

carácter de prototipo de la aplicación con el propósito de ser presentado frente a la municipalidad de la ciudad y las principales empresas turísticas de la localidad para su financiamiento.



Imagen 1. Pantalla principal de la app

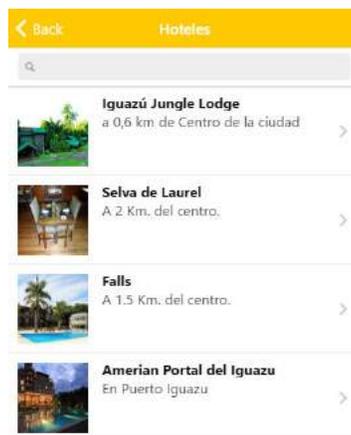


Imagen 2. Pantalla listado de hoteles.

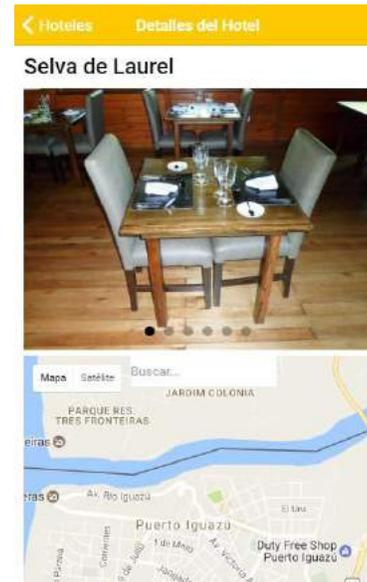


Imagen 3. Información de uno de los restaurantes seleccionado

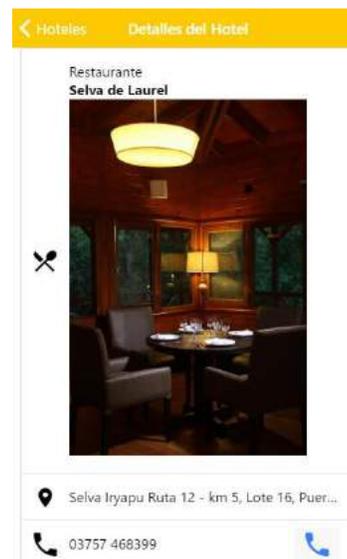


Imagen 4. Información de uno de los restaurantes seleccionado

Conclusión

El presente trabajo deja en evidencia el actual estado que se encuentra los servicios de informática, y las condiciones que se encuentran dadas para que cualquier persona pueda desplegar una idea que llegue muy lejos.

Con respecto a la tecnología presente hoy en día en cuanto al cloud computing es una tecnología que planteó un nuevo paradigma y que vino para quedarse, y será una verdadera revolución que se dará como se dio en su momento con el Internet y la Web.

El presente proyecto representa una gran innovación para la Ciudad de Puerto Iguazú, para ciudades aledañas y localidades en un futuro, debido al gran número de teléfonos inteligentes y su uso masivo por parte de los turistas.

Los turistas son grandes consumidores de aplicaciones móviles que les permitan mejorar su experiencia en el lugar que se encuentran visitando. La aplicación que se desarrolla con este proyecto juntamente con la incorporación de tecnologías como ser GeoMarketing, Geolocalización y el concepto de SoLoMo servirán para estudiar los hábitos de los turistas y mejorar la oferta turística actual y así ampliarla. Además, dará origen a nuevos emprendimiento a partir del análisis detallado de los datos que proporcione el uso de la aplicación y el estudio del consumo por parte de los turista; como por ejemplo a partir del análisis de flujo de personas por ciertos lugares, el recorrido que hacen los turistas dentro de la ciudad y caminos por donde se desplazan se pueden armar nuevos circuitos turísticos.

Todo lo anterior llevará a aumentar la tasa de permanencia del turista en la ciudad de Iguazú, propiciando así el aumento del consumo y el fomento de la región por parte del sector turístico.

Se debe aclarar que el hecho de ofrecer y promocionar puntos turísticos que no pertenecen a la ciudad de Puerto Iguazú, no representa una amenaza al cumplimiento del objetivo general del presente proyecto.

Referencias

COMPUTACIÓN EN LA NUBE (CLOUD COMPUTING). Diapositiva del curso de Sistemas Distribuidos de la Maestría en Tecnologías de la Información. Universidad

Nacional del Nordeste (UNNE) y la Universidad Nacional de Misiones (U.Na.M.). Consultado el 1 de diciembre de 2016.

L. Joyanes Aguilar. Cloud Computing. Estrategias digitales para organizaciones y empresas. AlfaOmega, México, 2011. Consultado el 1 de diciembre de 2016.

Cómputo Ubicuo e Interacción Humano-Computadora. Jesus Favela. CICESE. Material del curso de Sistemas Distribuidos de la Maestría en Tecnologías de la Información. Universidad Nacional del Nordeste (UNNE) y la Universidad Nacional de Misiones (U.Na.M.). Consultado el 1 de diciembre de 2016.

Novatica. Computación Ubicua. Material del curso de Sistemas Distribuidos de la Maestría en Tecnologías de la Información. Universidad Nacional del Nordeste (UNNE) y la Universidad Nacional de Misiones (U.Na.M.). de la cátedra. Consultado el 1 de diciembre de 2016.

Presentación. Desarrollo de Aplicaciones Móviles Sensibles al Contexto. Lic. en Cs. de la Comp. e Ingeniería en Computación. Depto. de Ciencias e Ingeniería de la Computación. Universidad Nacional del Sur. 1er. Cuatrimestre de 2016. Disponible en <http://cs.uns.edu.ar/~mvm/CAIA-2016/index.php?accion=download&dir=downloads/Clases> Consultado el 1 de diciembre de 2016.

La información científica en la UNdeC:

¿Aumento la visibilidad web?

Jose Texier¹, Fernando Emmanuel Frati¹, Fernanda Carmona¹, Alberto Riba¹, Emiliano Peressini¹, Jusmeidy Zambrano¹, Rita Tosetto¹, Antonio Castro Lechtaler²

¹Departamento de Ciencias Básicas y Tecnológicas, Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina

²CISTIC/FCE - Universidad de Buenos Aires

{jtexier, fefrati, fbcarmona, ariba}@undec.edu.ar, emiliano.peressini@gmail.com, jzambrano@undec.edu.ar, rtosetto@undec.edu.ar, acastro@est.iue.edu.ar

RESUMEN

En la actualidad se están reestructurando los procesos sociales gracias al desarrollo de Internet y de las Tecnologías de Información y Comunicación (TIC). De allí que el eje se ha volcado en lo cultural, social, económico y tecnológico a la información, como materia prima [4, 16]. De manera tal que existe una perenne necesidad de la gestión, difusión y preservación de esa información a través, por ejemplo, de plataformas de software como las Bibliotecas Digitales (BD) y los Repositorios Institucionales (RI), ya que ambos ofrecen servicios similares y el uso de cada uno de términos (RI y BD) depende del ámbito donde se aplica y, por ende, de los recursos con los que se desean trabajar [15]. En este sentido, la Universidad Nacional de Chilecito (UNdeC) se propone posicionar su producción científica a través de una estrategia de gestión de la información de forma eficaz y eficiente, maximizando su visibilidad y garantizando la generación de productos científicos de calidad, ya sea en Bibliotecas Digitales y/o Repositorios Institucionales. Con esta línea de I/D/I se está generando conciencia en la comunidad universitaria de la UNdeC

respecto de algunos términos de uso frecuente en las áreas de visibilidad y gestión del conocimiento.

Palabras clave: producción científica, repositorios institucionales, acceso abierto, recursos, visibilidad web.

CONTEXTO

La línea de I/D/I corresponde al diseño, desarrollo e implementación de proyectos que fortalecerán la investigación y las diferentes labores relacionadas con la gestión del conocimiento en cuanto a la visibilidad web y la preservación, de la mayoría, de la producción académica y científica de la UNdeC.

Actualmente se encuentra activo un PDTs (Proyectos de Desarrollo Tecnológico y Social) aprobado el año pasado y titulado “Repositorios digitales con contenidos orientados a las necesidades de escuelas rurales primarias y secundarias”. Su ejecución comenzó en abril del 2016 y finaliza en marzo del 2018. También se encuentra activo el proyecto “Red para la creación y publicación de objetos virtuales de aprendizajes de calidad en Repositorios Institucionales” correspondiente a la

convocatoria “Redes Internacionales 9”, promovida por la Secretaría de Políticas Universitarias de la Nación. De igual manera, se tiene un proyecto del 2013-2014 del programa “Financiamiento para el Estímulo y Desarrollo de la Investigación Científica y Tecnológica” de la Secretaría de Ciencia y Tecnología en la UNdeC (FICyT – UndeC). El título es “Fortalecimiento y visibilidad web de la producción científica de la UNdeC”, que finaliza en julio de este año.

En cuanto a la participación del grupo de trabajo en consorcios internacionales, se están desarrollando dos (2) propuestas para vincular la UNdeC con organizaciones internacionales: Latin American and Caribbean Consortium of Engineering Institution (LACCEI) e Ibero american Science and Technology Education Consortium (ISTEC-BIREDIAL).

Por otra parte, se destaca que la UNdeC cuenta con la estructura tecnológica y de RRHH necesarios para ejecutar los proyectos que surjan de la línea de investigación.

1. INTRODUCCIÓN

Gestión/organización del conocimiento es un concepto que ha ganado cierta difusión. Barité y Dahlberg [1], [2] señalan que cuando se habla de *organización del conocimiento* “es importante tener presente que (...) se hace referencia (...) [al] conocimiento socializado, compartido o comunicado, que además ha sido registrado”.

Por tanto, la información es la base para acceder al conocimiento. Se entiende al conocimiento como "una entidad abstracta que existe solamente en la mente de un ser humano en tanto sujeto

cognoscente (es decir, es lo que yo sé), mientras que la información es el conocimiento comunicado, compartido o socializado" [3]. Esta diferenciación entre ambos conceptos (conocimiento e información) nos permiten comprender por qué es necesario que la información esté organizada para que el acceso y la apropiación [14] por parte de los sujetos les haga partícipes de la sociedad y los empodere como ciudadanos informados. Lo que en palabras de Barité y Dahlberg apunta a la socialización del conocimiento.

Sin embargo, esta situación se hace posible si esa información puede (o debe) registrarse en un soporte físico para convertirse en un documento o recurso bibliográfico con genuina utilidad por quien lo desee, por lo que los Repositorios Institucionales (RI) bajo la filosofía del Acceso Abierto pueden convertirse en una de las vías más idóneas para lograr la visibilidad web que requiere la Institución.

Se entiende que los **Repositorios Institucionales (RI)** están constituidos por un conjunto de archivos digitales en representación de productos científicos y académicos que pueden ser accedidos por los usuarios [4]. En otras palabras, los RI son estructuras web interoperables que alojan recursos científicos, académicos y administrativos, tanto físicos como digitales, descritos por medio de un conjunto de datos específicos (metadatos) [5]–[7]. Tienen como propósito recopilar, catalogar, gestionar, acceder, difundir y preservar la información [4], [8]. Vale la pena destacar que los RI son vías de comunicación científica, pero no son canales de publicación. Eso quiere decir que se deben seguir los mismos mecanismos de validación científica

existentes hasta ahora a través de las revisiones por pares, pero los autores deben hacer énfasis en mostrar sus publicaciones y datos primarios de sus investigaciones. Se destaca que los RI deben disponer también de Objetos de Aprendizaje (OA), que son entidades generalmente entregadas a través de Internet y diseñadas con el fin de que sean utilizadas y reutilizadas en múltiples contextos educativos [12, 13].

Los RI se configuran dentro de la filosofía del **Acceso Abierto** (en inglés *Open Access - OA*). Esta filosofía tiene como fin asegurar el acceso libre y abierto a la producción científica, es decir, garantizar el acceso a través de Internet sin que los derechos del copyright sean una barrera [9], [10]. Los RI materializan el objetivo del OA porque la información que se deposita es una producción que tienen como propósito ser accesible, sin restricciones y preservada digitalmente.

Sobre la base de lo expuesto, esta línea de investigación desarrollará proyectos que fortalezcan la difusión y gestión de la comunicación científica y académica de los becarios, docentes e investigadores de cualquier disciplina de la UNdeC. Esto se logrará mediante la socialización y sistematización de un conjunto de conocimientos prácticos que permitan alcanzar una visibilidad web que ayude a mejorar el posicionamiento de la UNdeC en el país y en el mundo. Todo esto bajo el contexto de la filosofía del Acceso Abierto.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

- Repositorios institucionales y bibliotecas digitales.

- Objetos de aprendizaje.
- Gestión de la información y el conocimiento.
- Sistemas de información web y bases de datos.
- Índices bibliométricos.
- Interoperabilidad.
- Preservación digital.
- Recuperación de la información.
- Análisis semántico de la información científica.

3. OBJETIVOS

El desarrollo de esta línea en la UNdeC esta permitiendo cumplir con los siguientes objetivos:

- Difundir el concepto de conocimiento y la distinción de dato e información.
- Comprender los problemas y las tendencias asociadas a la producción abierta del conocimiento, el uso y la difusión.
- Analizar la visibilidad web de la UNdeC.
- Conocer la mayoría de la producción científica de la UNdeC.
- Analizar los recursos educativos existentes que cumplan con las necesidades educativas de la UNdeC y la ciudad de Chilecito.
- Fomentar la formación de habilidades y conocimientos relacionados con los RI.
- Desarrollar un mapa conceptual de la producción científica de la UNdeC.
- Entender y valorar el movimiento mundial de Acceso Abierto y señalar las vías de su materialización.

- Desarrollar las estrategias necesarias para optimizar la visibilidad científica de la UNdeC.
- Implementar Repositorios de prueba para visualizar los diferentes recursos (educativos, académicos y científicos) que se produzcan en Chilecito.

4. RESULTADOS

OBTENIDOS/ESPERADOS

- Conocimiento de la producción científica, lugar(es) donde está alojada y qué personal la realizó. En julio del 2016 se realizó un relevamiento de los productos científicos con filiación UNdeC, de esta manera, se obtuvo una ponderación del posicionamiento web de la UNdeC.
- Relevamiento de los diferentes recursos educativos producidos en la UNdeC y en la ciudad de Chilecito, sobre los linemamientos del proyecto PDTS, anteriormente nombrando.
- Elaboración de un mapa conceptual de la producción científica.
- Curso a nivel de postgrado llamado “Gestión de la Información Académica y Científica”.
- Diseño de un conjunto de estrategias para optimizar la producción científica de la UNdeC.
- Se desarrolló un proyecto de asignaturas sobre las Bibliotecas Digitales para presentarse en dos universidades [15, 16].
- Se elaboró un software que permitió la extracción de metadatos de artículos de diferentes fuentes para

normalizarla y visualizarla [17, 18].

- Se implementó un repositorio de prueba que esta gestionando los recursos educativos y objetos de aprendizaje generados dentro del proyecto PDTS.
- Se están elaborando dos artículos, uno sobre la producción científica de la UNdeC y otro sobre un analizador semántico de los artículos científicos.

5. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por cuatro docentes de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas de la UNdeC (acreditadas por CONEAU), dos doctores especializados en repositorios institucionales, bibliotecas digitales, desarrollo de software, cómputo paralelo y tecnología grid. Otra docente que está finalizando su doctorado y dos definiendo su tesis de Maestría en Informática. También participa un alumno avanzado de grado. En otras palabras, se cuenta con un recurso humano con habilidades y formación académica en las diversas áreas de la propuesta, asegurando la concreción de la línea, por ejemplo, uno es especialista en RI y OA, con investigación en esa área y otra esta en formación (maestría) sobre el área de los Objetos de Aprendizaje. Adicionalmente, se destaca que dos están categorizados en el programa de incentivos.

Los integrantes son docentes de las asignaturas Programación I, Sistemas I, Arquitecturas Paralelas, Teoría de la Computación y Herramientas de Ingeniería de Software. Estas asignaturas contemplan la aprobación mediante la

participación en proyectos de investigación, por lo que pueden surgir nuevos trabajos en esta línea.

6. BIBLIOGRAFÍA

- [1] M. Barité, *Diccionario de organización y representación del conocimiento*. Montevideo, 2000.
- [2] I. Dahlberg, “Knowledge organization: its scope and possibilities,” *Knowledge organization*, vol. 20, no. 4.
- [3] A. M. Martínez-Tamayo and J. C. Valdez, *Indicación y Clasificación en Bibliotecas*, Primera edición. Argentina: Alfagrama, 2009.
- [4] J. Texier, “Los repositorios institucionales y las bibliotecas digitales: una somera revisión bibliográfica y su relación en la educación superior,” presented at the 11th Latin American and Caribbean Conference for Engineering and Technology - 2013, Cancun, Mexico, 2013, p. 9.
- [5] C. A. Lynch, “Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age,” *ARL*., 2003. [Online]. Available: <http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>. [Accessed: 28-Oct-2013].
- [6] J. Tramullas and P. Garrido, “Software libre para repositorios institucionales: propuestas para un modelo de evaluación de prestaciones,” *El Profesional de la Información*, vol. 15, no. 3, pp. 171–181, 2006.
- [7] H. Van de Sompel, S. Payette, J. Erickson, C. Lagoze, and S. Warner, “Rethinking Scholarly Communication,” *D-Lib Magazine*, vol. 10, no. 9, Sep. 2004.
- [8] M. De Giusti, N. Oviedo, A. Lira, A. Sobrado, J. Martinez, and A. Pinto, “SEDICI – Desafíos y experiencias en la vida de un repositorio digital,” *RENATA*, vol. 1, no. 2, pp. 16–33, Aug. 2011.
- [9] P. Suber, “Ensuring open access for publicly funded research,” *BMJ*, vol. 345, 2012.
- [10] D. Torres-Salinas, N. Robinson-García, and A. Cabezas-Clavijo, “Compartir los datos de investigación en ciencia: introducción al data sharing,” *Profesional de la Información*, vol. 21, no. 2, pp. 173–184, 2012.
- [11] P. Suber, “Open access, impact, and demand,” *BMJ*, vol. 330, no. 7500, pp. 1097–1098, May 2005.
- [12] D. Wiley, “Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy,” presented at the Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos, SPEDECE, 2001.
- [13] E. Morales, F. Garcia, A. Barron, A. Berlanga, and C. Lopez, “Propuesta de Evaluación de Objetos de Aprendizaje,” presented at the Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos Reutilizables (SPDECE), Barcelona, 2005.
- [14] Kalman J. (2001). *Saber lo que es la Letra*. Informe. México: DIE-CINVESTAV-IPN.
- [15] Texier, J., Zambrano, J., & Carmona, F. B. (2016). *Las Bibliotecas Digitales en el Currículum de las carreras de Ciencias de la Computación: una propuesta posible para Argentina y Venezuela*. In XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016).

- [16] Texier, J. (2016). Los repositorios institucionales y su importancia en las Universidades Venezolanas. *Revista Scitus*. Venezuela.
- [17] Villarreal, G. L., Terruzzi, F. A., De Giusti, M. R., Lira, A. J., & Texier, J. D. (2016). Fostering the institutional repository through policies and interoperability with online services: the case of Universidad Nacional de La Plata. *Scholarly and Research Communication*, 6(1).
- [18] Texier, J., Zambrano, J., & Riba, A. (2016). Normalización de los LACCEI Proceedings a través de un proceso ETL.

USO DE VRPN EN LA IMPLEMENTACIÓN DE UNA BCI PARA REHABILITACIÓN NEUROLÓGICA

Javier J. Rosenstein^{1,2,3}, Osvaldo E. Marianetti², Raúl E. Otoya Bet³

¹Instituto de Investigaciones, Facultad de Informática y Diseño, Universidad Champagnat, Belgrano 721, 5501 Godoy Cruz, Mendoza, Mendoza, Argentina.
rosensteinjavier@uch.edu.ar

²Universidad de Mendoza, Dirección de Posgrado, Facultad de Ingeniería, Boulogne Sur Mer 683, 5500 Mendoza, Argentina.
omarianetti@um.edu.ar

³NNT Neurotechnology, Neuromed Argentina S.A., Leguizamón 398, 5501 Mendoza, Argentina.
{rosenstein.javier},{otoya.raul}@neuromed.com.ar

RESUMEN

En el desarrollo de sistemas de realidad virtual uno de los inconvenientes que se encuentran es la comunicación entre las aplicaciones y los dispositivos de adquisición, ya sea por no disponer de un método de acceso en forma directa de los dispositivos o por necesitar independencia entre ambos, es decir que las aplicaciones corran en una plataforma y los dispositivos en otras. Para lograr esta independencia y a su vez permitir la integración de todo el sistema de realidad virtual, es necesario la implementación de algún protocolo de comunicaciones que permita esta vinculación heterogénea en tiempo real. Los dispositivos generalmente están asociados a funciones o características de los individuos que los utilizan y se necesita integrar los movimientos que estos representan a la aplicación de Realidad Virtual correspondiente. El presente trabajo trata del análisis e implementación del protocolo de comunicaciones VRPN (Virtual Reality Protocol Network) entre las partes de un entorno multimedia donde interactúan la adquisición de movimientos del usuario y la representación Visual en un escenario virtual

que permita la retroalimentación al usuario en tiempo real logrando una experiencia interactiva e inmersiva, Esto tiene aplicación directa en los tratamientos de rehabilitación en pacientes de patologías neurológicas y cognitivas.

Palabras clave: VRPN, EOG, EEG, BCI, Serious Games, Mirror Neurons, Neurofeedback

CONTEXTO

El presente proyecto forma parte del trabajo de tesis de Maestría en Teleinformática de la Universidad de Mendoza (Ciudad, Mendoza), correspondiente al tesista Javier Rosenstein, el mismo se desarrolla en el Instituto de Investigaciones de la Facultad de Informática y Diseño de la Universidad Champagnat (Godoy Cruz, Mendoza), en el marco de la Licenciatura en Sistemas de Información; en cooperación con el Laboratorio de I+D+i en Neurotecnologías de la empresa Neuromed Argentina S.A. (Godoy Cruz, Mendoza).

Este trabajo es parte del proyecto de investigación que dio inicio en Diciembre de

2017 denominado “Diseño y desarrollo de un prototipo de Serious Game destinado a la rehabilitación de problemas neurológicos implementando VRPN para la comunicación de la BCI”.

1. INTRODUCCIÓN

El objetivo principal de la presente línea de investigación consiste en el diseño y desarrollo de una BCI (Brain Computer Interface) [1], [2], que permita interactuar entre las señales generadas por un paciente neurológico [3] y luego de ser interpretadas las mismas poder ser enviadas hacia una interfaz de realidad virtual la cual nos permita lograr el principio de neurofeedback [4], retroalimentación hacia el paciente y de este modo mejorar sus capacidades cognitivas correspondientes, e incluso estudios indican la posibilidad de tratar patologías del comportamiento psiquiátricas como la depresión [5] y [6]. La implementación del trabajo se organiza de acuerdo a las siguientes etapas:

- a) Adquisición de señales mediante técnicas de Electroencefalografía (EEG) [7] y adquisición de movimientos oculares mediante las técnicas de electrooculografía (EOG) [8], [9], [10], [11] y [12].
- b) Análisis rápido de éstas señales para poder identificar la voluntad de movimiento del individuo, así como la dirección del movimiento.
- c) Codificar lo anterior a comandos en el protocolo VRPN (Virtual Reality Protocol Network) que permita transmitir la información al componente software/hardware que la requiera.
- d) Dirigir una Interfaz gráfica de aprendizaje o Interfaz Cerebro/Computadora (BCI – Brain Computer Interface).

Finalmente se presenta la BCI como un sistema de adquisición de datos,

procesamiento del protocolo serie a VRPN y luego la representación en nuestro modelo de prototipo de Serious Game. [14] y [15]. Una vez adquiridas las señales EEG/EOG éstas se analizan y codifican con las librerías cliente desarrolladas como parte de este proyecto, lo que permite su comunicación mediante todo el sistema VRPN hacia las interfaces virtuales que interpretan este protocolo.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto está compuesto por tres etapas o fases de trabajo. La primera de ellas corresponde a la implementación de VRPN en un sistema de captura de datos simulada de un usuario, procesamiento y comunicación de estos hacia una interface virtual básica, La segunda etapa del proyecto pretende avanzar sobre la captura de datos reales del usuario ya no simuladas para que luego de procesadas ingresen al sistema BCI y utilicen todo lo obtenido por la primera etapa, es decir la implementación de VRPN y la interface virtual de neurofeedback. Una breve descripción de cada una de las etapas se describe a continuación:

1. La primera parte consiste en la implementación de un simulador de señales de EEG y EOG necesarias para el análisis e interpretación de la voluntad del usuario de la BCI, estas señales una vez procesadas se deben codificar en comandos de VRPN para poder ser transmitidas hacia una interface virtual, esta debe poder interpretar las señales transmitidas y representar la voluntad inicial del usuario correspondiente. De este modo se cumplen los objetivos de captura, análisis, procesamiento, transmisión, recepción y representación lo cual produce el efecto de neurofeedback deseado sobre una

- interface virtual de capacitación a nivel prototipo.
2. La segunda etapa consiste en el desarrollo y mejora del proyecto mediante la adquisición real de las señales de EEG y EOG por electrodos ubicados superficialmente sobre la cabeza del usuario o paciente, esta modificación al sistema requiere de un diseño e implementación electrónica así como del desarrollo del firmware que permita adquirir, analizar y preprocesar estas señales para ser transmitidas hacia el equipo que genera a partir de la recepción de estas las tramas de acuerdo al protocolo VRPN y así se incorpora en forma transparente al proyecto implementado en la primer etapa correspondiente, luego de esta etapa el proyecto BCI está completo.
 3. La etapa final de este proyecto consiste en el diseño y desarrollo del escenario virtual de rehabilitación cognitiva para pacientes neurológicos según las indicaciones concretas por parte del especialista en neurología, partiendo de un relevamiento de las técnicas de aprendizaje que se requieren implementar y los resultados que se pretenden obtener, indicados por el neurólogo o experto afín.

3. RESULTADOS ESPERADOS

Como resultados esperados, los mismos pueden dividirse en dos grandes grupos:

1. *BCI e interface virtual de neurofeedback*, se espera obtener como producto final un sistema de retroalimentación a un supuesto usuario simulado por la generación de señales que comandan el escenario de realidad virtual via comandos codificados en el protocolo VRPN que viajan via una red ethernet.

2. *Adquisición de datos reales de EEG y EOG, su analisis y preprocesamiento identificando los comandos necesarios para su transmisión hacia el equipo generador de tramas VRPN para que finalmente se tenga el proyecto totalmente terminado y funcional que implemente el efecto de neurofeedback propuesto.*

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de I+D presentada está vinculada con el desarrollo de una tesis de postgrado, por parte del estudiante de maestría en teleinformática de la Universidad de Mendoza, Javier J. Rosenstein. Dicha tesis se centra en la implementación del protocolo VRPN demostrando su uso en un sistema BCI.

Además una vez que el proyecto se encuentre implementado desde su primer etapa será utilizado como recurso para el dictado de talleres de comunicaciones, redes, programación de microcontroladores, Programación en C/C++ y Diseño y programación de interfaces virtuales de capacitación/rehabilitación en general, tanto para estudiantes de la universidad, como así también para aquellos interesados externos a la UCH.

5. BIBLIOGRAFÍA

1. 12. Y. Wang, X. Gao, B. Hong, and S. Gao, "Practical designs of brain-computer interfaces based on the modulation of eeg rhythms," in *Brain-Computer Interfaces*. plus 0.5em minus 0.4emSpringer, 2009, pp. 137–154.
2. 15. J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control,"

- Clinical neurophysiology, vol. 113, no. 6, pp. 767–791, 2002.
3. J. A. Pineda, “The functional significance of mu rhythms: translating “seeing” and “hearing” into “doing”,” *Brain Research Reviews*, vol. 50, no. 1, pp. 57–68, 2005.
 4. S. Enriquez-Geppert, R. J. Huster, and C. S. Herrmann, “Boosting brain functions: Improving executive functions with behavioral training, neurostimulation, and neurofeedback,” *International Journal of Psychophysiology*, vol. 88, no. 1, pp. 1–16, 2013.
 5. R. Ramirez, M. Palencia-Lefler, S. Giraldo, and Z. Vamvakousis, “Musical neurofeedback for treating depression in elderly people.” *Frontiers in neuroscience*, vol. 9, pp. 354–354, 2014.
 6. W. Rief, “Getting started with neurofeedback,” 2006.
 7. J. D. Kropotov, *Quantitative EEG, event-related potentials and neurotherapy*. Academic Press, 2010.
 8. A. Bulling, J. A. Ward, H. Gellersen, and G. Troster, “Eye movement analysis for activity recognition using electrooculography,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 4, pp. 741–753, 2011.
 9. H. Singh and J. Singh, “A review on electrooculography,” *International Journal of Advanced Engineering Technology*, vol. 3, no. 4, pp. 115–122, 2012.
 10. D. P. Bautista, I. A. Badillo, D. De la Rosa Mejía, and A. H. H. Jiménez, “Interfaz humano-computadora basada en señales de electrooculografía para personas con discapacidad motriz,” *ReCIBE*, vol. 3, no. 2, 2016.
 11. S. Yathunathan, L. Chandrasena, A. Umakanthan, V. Vasuki, and S. Munasinghe, “Controlling a wheelchair by use of eog signal,” in 2008 4th International Conference on Information and Automation for Sustainability. plus 0.5em minus 0.4emIEEE, 2008, pp. 283–288.
 12. V. C. C. Roza, “Interface para tecnologia assistiva baseada em electrooculografia,” 2014.
 13. A. C. Gaviria, I. C. Miller, S. O. Medina, and D. R. Gonzales, “Implementación de una interfaz hombre-computador basada en registros eog mediante circuitos de señal mixta psoc,” in V Latin American Congress on Biomedical Engineering CLAIB 2011 May 16-21, 2011, Habana, Cuba. plus 0.5em minus 0.4emSpringer, 2013, pp. 1194–1197.
 14. P. Rego, P. M. Moreira, and L. P. Reis, “Serious games for rehabilitation: A survey and a classification towards a taxonomy,” in 5th Iberian Conference on Information Systems and Technologies. plus 0.5em minus 0.4emIEEE, 2010, pp. 1–6.
 15. J. S. Breuer and G. Bente, “Why so serious? on the relation of serious games and learning,” *Eludamos. Journal for Computer Game Culture*, vol. 4, no. 1, pp. 7–24, 2010.

Una Herramienta Gráfica con Razonamiento basado en DL para el Análisis de Modelos OVM

Angela Oyarzun¹

Germán Braun^{1,2,3}

Laura Cecchi¹

Pablo Fillottrani^{2,4}

email: angela.oyarzun@est.fi.uncoma.edu.ar, {german.braun,lcecchi}@fi.uncoma.edu.ar,
prf@cs.uns.edu.ar

¹*Grupo de Investigación en Lenguajes e Inteligencia Artificial*

Departamento de Teoría de la Computación - Facultad de Informática

UNIVERSIDAD NACIONAL DEL COMAHUE

²*Laboratorio de I&D en Ingeniería de Software y Sistemas de Información*

Departamento de Ciencias e Ingeniería de la Computación

UNIVERSIDAD NACIONAL DEL SUR

³*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*

⁴*Comisión de Investigaciones Científicas de la provincia de Buenos Aires (CIC)*

Resumen

Esta línea de investigación se desarrolla en forma colaborativa entre docentes investigadores de la Universidad Nacional del Comahue y de la Universidad Nacional del Sur, en el marco de proyectos de investigación financiados por las universidades antes mencionadas.

El objetivo general del trabajo de investigación es el desarrollo de una herramienta Web que permita integrar el soporte gráfico para el diseño de modelos de variabilidad ortogonal (OVM) y el razonamiento automático para validar dichos modelos. Se trabajará en una arquitectura cliente-servidor, en la definición de un entorno gráfico con primitivas basadas en OVM y en la traducción de dichos modelos a la lógica descriptiva *ALCCZ*. De esta manera, el usuario podrá diseñar y visualizar modelos OVM y, a su vez, analizar la consistencia de los mismos.

Palabras Clave: Modelos de Variabilidad Ortogonal, Lógicas Descriptivas, Ingeniería de Software basada en Conocimiento

Contexto

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el marco del proyecto de investigación *Agentes Inteligentes y Web Semántica (04/F014)* y a través de una Beca de Iniciación a la Investigación para Alumnos, por la Universidad Nacional del Sur a través del proyecto de investigación *Integración de Información y Servicios en la Web (24/N027)* y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral. Los proyectos de investigación tienen una duración de cuatro años y la beca doctoral una duración de 5 años, finalizando esta última en abril de 2019. La beca para alumnos de iniciación a la investigación tiene una duración de un año.

1. Introducción

El modelo de variabilidad ortogonal (OVM) [1] se encarga de definir la variabilidad de una línea de productos de software (SPL), relacionándola con otros modelos como los casos de usos, diagramas de clases, modelos de características (FM), entre otros. Estos modelos OVM son utilizados durante la fase de ingeniería de dominio de una SPL como parte del proceso de gestión de variabilidad, el cual abarca la definición, modelado, implementación y validación de las características variables de los productos software derivados de una línea. La gestión de variabilidad está incluida tanto en la fase de ingeniería de dominio como en la de ingeniería de aplicación del desarrollo de una SPL y es una de las actividades más importantes ya que ésta se ve reflejada en la manera en que el software es desarrollado, extendido y mantenido.

Dado el gran impacto de los modelos OVM sobre la calidad de los productos, surge el análisis de variabilidad automático, con el objeto de detectar errores en etapas tempranas de desarrollo. Este análisis [2, 3, 4, 5] se enfoca en un conjunto de técnicas para traducir y validar los modelos de variabilidad considerando las anomalías o incompatibilidades que pueden contener. En este sentido, chequear la consistencia de los modelos de variabilidad es un problema crítico.

Actualmente, existen trabajos relacionados con el análisis de variabilidad automatizado que proponen diferentes técnicas y métodos [5, 6]. Entre ellos se encuentra, FaMa-OVM [6], una herramienta que trabaja con modelos OVM descritos mediante un formato textual, los cuales son posteriormente analizados utilizando SAT-solvers. Esta herramienta, al igual que las diferentes propuestas basadas en SAT, presentan ciertas limitaciones ligadas a las lógicas restrictivas, razón por la cual no logran reflejar de manera correcta la estructura lógica de los modelos de variabilidad.

En este sentido, nuestro grupo presentó una propuesta [7] que se enfoca en la utili-

zación de lógicas descriptivas para mejorar el análisis automatizado. En el trabajo mencionado, se formaliza en lógica descriptiva *ALCCT* el framework SeVaTax[8], basado en extensiones de OVM y cuyo objetivo es el análisis de las propiedades de los modelos de variabilidad y la derivación de productos a partir de una SPL. Asimismo, se estableció que el razonamiento sobre SeVaTax está en EXPTIME.

En el ámbito de este trabajo, proponemos desarrollar una herramienta Web que permita el diseño y la visualización de los diagramas de SeVaTax, ampliando el soporte gráfico fuertemente integrado con el razonamiento automático. De esta manera, se busca evitar la ocurrencia de errores provocados en el diseño de los modelos. Para lograr esto, se traducirán los diagramas SeVaTax a *ALCCT* según la formalización propuesta en [7] a fin de poder analizarlos de manera más precisa, utilizando razonadores lógicos basados en DL [9] y protocolos de comunicación como OWLlink [10]. OWLlink es una evolución de DIG para el lenguaje OWL 2 [11].

La estructura del presente trabajo es la siguiente. En la sección 2 presentamos los objetivos de los proyectos de investigación en los que se enmarca este trabajo y describimos la línea de investigación actual. En la sección 3 indicamos algunos resultados obtenidos y trabajos futuros. Finalmente, comentamos aspectos referentes a la formación de recursos humanos en esta temática.

2. Línea de Investigación y Desarrollo

Por un lado, en el proyecto de investigación *Agentes Inteligentes y Web Semántica*, se investigan diversas técnicas de representación de conocimiento y razonamiento, metodologías de modelado conceptual y mecanismos para la interoperabilidad de aplicaciones, y se hace énfasis en la aplicación de estos conceptos como soporte para desarrolladores de ontologías.

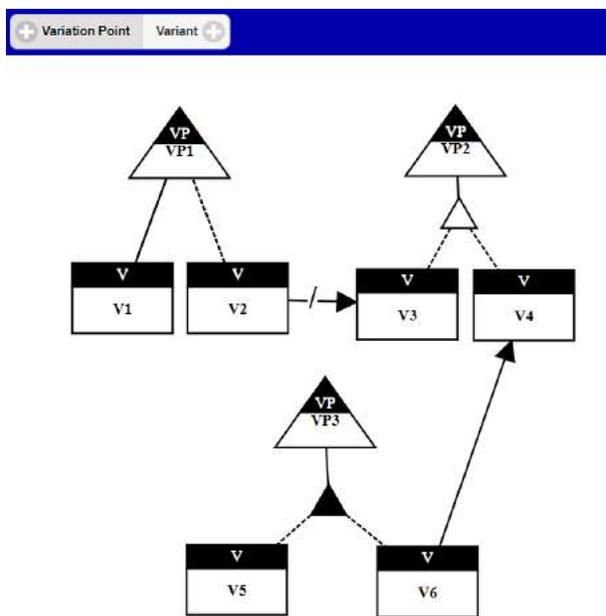


Figura 1: Captura de pantalla de la interfaz Web en desarrollo.

Por otro lado, en el proyecto de investigación *Integración de Información y Servicios en la Web* se propone investigar y desarrollar metodologías y herramientas que favorezcan la interoperabilidad semántica de información y de servicios en la Web, fundamentados en los últimos avances en el área de lenguajes de representación del conocimiento, ontologías y modelado conceptual.

Ambos proyectos confluyen en la línea de investigación de este trabajo, en la que se explora entre otros, sobre temas afines a la Representación del Conocimiento, las Lógicas Descriptivas [12], las Ontologías, la Ingeniería de Software basada en Conocimiento y la Ingeniería de Conocimiento. Particularmente, se ha escogido experimentar sobre metodologías que integren razonamiento con un *front-end* gráfico para dar soporte a la ingeniería de ontologías.

En esta línea de investigación se propone como principal objetivo desarrollar una herramienta Web basada en *crowd* [13, 14]. *crowd* es una plataforma Web con soporte gráfico para el diseño de modelos conceptuales como EER, UML y ORM2, integrada a

sistemas de razonamiento que permiten validarlos. Sin embargo, a diferencia de *crowd*, esta nueva herramienta sólo se centrará en los modelos de variabilidad ortogonal y en particular en los diagramas de SeVaTax.

La herramienta tendrá una arquitectura cliente-servidor. El lado del cliente constará de una interfaz de usuario que ofrecerá un conjunto de primitivas gráficas propias del lenguaje OVM extendido en SeVaTax, con las cuales el usuario podrá construir sus modelos. Mientras que el lado del servidor soportará un módulo traductor, el cual se encargará de codificar los diagramas de SeVaTax en la lógica descriptiva *ALCC* para, finalmente, ser enviado al razonador como un mensaje OWLink.

3. Resultados Obtenidos y Trabajo Futuro

A partir del análisis de la arquitectura de *crowd* se diseñó la arquitectura cliente - servidor de la herramienta, que incluye entre otros los módulos gráfico y traductor para OWLink.

En este momento, el trabajo se ha centrado en el módulo gráfico, con el objeto de hacer posible que el usuario visualice y edite modelos de variabilidad ortogonal a través de una interfaz Web. En la Figura 1 se muestra una captura de pantalla de la interfaz Web de la herramienta en desarrollo.

En el módulo gráfico deben estar definidas las primitivas gráficas del lenguaje OVM. Para ello se optó por hacer uso de la biblioteca JointJS¹, también empleada por la herramienta *crowd*, dada la gran variedad de funcionalidades que ofrece, la utilización de Backbone² y la capacidad que brinda al poder expandir sus funciones creando y/o agregando plugins.

En este sentido, actualmente, se están utilizando los plugins ya disponibles, que proveen las primitivas para EER y UML, para

¹<http://www.jointjs.com/>

²<http://backbonejs.org>

construir uno nuevo donde estarán definidas cada una de las primitivas gráficas del lenguaje de modelado de variabilidad ortogonal.

Finalmente, como trabajo futuro se encuentra la implementación de las funcionalidades necesarias desde el front-end para permitir tanto la visualización como la interacción con los modelos. Asimismo, se trabajará en el módulo traductor para OWLink, que nos permitirá la comunicación con razonadores, y concretar un primer prototipo de la herramienta.

4. Formación de Recursos Humanos

Uno de los autores de este trabajo es becario doctoral CONICET y está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur.

En la Universidad Nacional del Comahue, Facultad de Informática, se otorgaron Becas de Iniciación a la Investigación para alumnos para estimular la vocación científica. Una de esas becas fue otorgada a uno de los autores de este trabajo, que desarrollará su tesis de grado de la carrera Licenciatura en Ciencias de la Computación en esta temática.

Referencias

- [1] Klaus Pohl, Günter Böckle, and Frank J. van der Linden. *Software Product Line Engineering: Foundations, Principles and Techniques*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.
- [2] Fabricia Roos-Frantz, José A Galindo, David Benavides, Antonio Ruiz Cortés, and J García-Galán. Automated analysis of diverse variability models with tool support. *Jornadas de Ingeniería del Software y de Bases de Datos (JISBD 2014)*, Cádiz, Spain, page 160, 2014.
- [3] David Benavides, Sergio Segura, and Antonio Ruiz-Cortés. Automated analysis of feature models 20 years later: A literature review. *Inf. Syst.*, 35(6):615–636, September 2010.
- [4] Matthias Kowal, Sofia Ananieva, and Thomas Thüm. Explaining anomalies in feature models. In *Proceedings of the 2016 ACM SIGPLAN International Conference on Generative Programming: Concepts and Experiences*, GP-CE 2016, pages 132–143, New York, NY, USA, 2016. ACM.
- [5] A. Metzger, K. Pohl, P. Heymans, P. Y. Schobbens, and G. Saval. Disambiguating the documentation of variability in software product lines: A separation of concerns, formalization and automated analysis. In *15th IEEE International Requirements Engineering Conference (RE 2007)*, pages 243–253, Oct 2007.
- [6] Fabricia Roos Frantz, José Ángel Galindo Duarte, David Felipe Benavides Cuevas, and Antonio Ruiz Cortés. FaMA-OVM: A tool for the automated analysis of ovms. In *Proceedings of the 16th International Software Product Line Conference - Volume 2. SPLC '12*, ACM, New York, NY, USA, 2012.
- [7] Germán Braun, Matias Pol'la, Laura Cecchi, Agustina Buccella, Pablo Fillottrani, and Alejandra Cechich. A DL Semantics for Reasoning over OVM-based Variability Models. 2017.
- [8] M. Pol'la, A. Buccella, M. Arias, and A. Cechich. Sevatax: service taxonomy selection validation process for spl development. In *2015 34th International Conference of the Chilean Computer Science Society (SCCC)*, pages 1–6, Nov 2015.
- [9] Volker Haarslev, Kay Hidde, Ralf Möller, and Michael Wessel. The racerpro knowledge representation and reasoning system. *Semantic Web Journal*, 2012.

- [10] Thorsten Liebig, Marko Luther, Olaf Noppens, and Michael Wessel. Owillink. *Semantic Web*, 2(1):23–32, 2011.
- [11] Markus Krötzsch. Owl 2 profiles: An introduction to lightweight ontology languages. In *Reasoning Web*, pages 112–183, 2012.
- [12] Diego Calvanese, Maurizio Lenzerini, and Daniele Nardi. Description logics for conceptual data modeling. In *Logics for Databases and Information Systems*, pages 229–263. Kluwer, 1998.
- [13] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Fillottrani. Una arquitectura cliente-servidor para modelado conceptual asistido por razonamiento automático. In *XVIII Workshop de Investigadores en Ciencias de la Computación*, 2016.
- [14] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Fillottrani. crowd: A tool for conceptual modelling assisted by automated reasoning - preliminary report. In *the 2nd Simposio Argentino de Ontologías y sus Aplicaciones SAOA '16 JAIIO '16 - to appear*, 2016.

Influencias del estado biométrico-emocional de personas interactuando en contextos de entornos virtuales

Jorge Ierache¹, Claudio Cervino², Edgardo Eszter³, Daniel Fortin³, Alvaro Castro Menna³

¹ISIER-UM -Facultad de Informática Ciencias de la Comunicación y Técnicas Especiales

²Instituto de Fisiología y Neurociencias (IFiNe- Secretaria de CyT, UM)

³Facultad de Ingeniería

Universidad de Morón

Cabildo 134, Buenos Aires, Argentina

jierache@unimoron.edu.ar

Resumen

Existen diferentes maneras de inferir emociones en los usuarios de computadoras, por ejemplo, a través de detección gestual de rostro, voz, registros de EEG, este último a través de *Brain Control Interface* (BCI). Actualmente se experimenta con el control de computadoras, dispositivos. Sin embargo no se registran desarrollos de sistemas que asistan a la inferencia emocional con capacidades de explotación y desarrollo de integración abierta y transparente a los entornos virtuales en especial de simulación y entrenamiento. Analizar las emociones y comportamiento de las personas en entornos virtuales que simulan de modo realista situaciones de la vida cotidiana o tareas laborales específicas, de riesgo, brindan un marco apropiado para, no sólo detectar, medir y analizar dichas emociones o respuestas de comportamiento, sino también para poder proyectarlas a situaciones ya no simuladas sino reales. Allí radica el valor de los simuladores como entorno para ensayos de esta naturaleza.

Palabras clave: Interfase cerebro-computadora, Reconocimiento patrones emocionales, Estado biométrico-

emocional, Entornos virtuales.

Contexto

Para el desarrollo de este proyecto se aúnan esfuerzos entre tres grupos de trabajo: el Instituto de Sistemas Inteligentes y Enseñanza Experimental de la Robótica (ISIER-UM), docentes-investigadores de la Facultad de Informática; Institutos de la Facultad de Ingeniería, docentes-investigadores de la Facultad de Ingeniería, e Instituto de Fisiología y Neurociencias (IFiNe) de la Secretaria de CyT-UM. Esta alianza, si bien de reciente conformación, está formado por docentes con una amplia trayectoria académica y de investigación en la UM, en el campo de la informática, ingeniería electrónica, la fisiología y las neurociencias. Este proyecto es financiado a través del Ping/17-03-JI-004 de la Secretaria de Ciencia y Tecnología de la Universidad de Morón.

Introducción

Durante las pasadas décadas se incrementó el desarrollo del campo de las Interfaces Cerebro-Maquina [1], [2] o más comúnmente conocidas por sus siglas en inglés de *Brain Control Interface*

(BCI). Esto posibilita la comunicación entre las funciones mentales y cognitivas de quien la utiliza para luego ser procesadas, clasificadas e interpretadas por aplicaciones o dispositivos puntuales. La investigación de las interfases BCI se desarrolla en un campo científico multidisciplinario con aplicaciones que van desde la computación aplicada en el campo de las neurociencias, domótica, robótica y entretenimiento entre otras; con aplicaciones que van desde mover cursores en pantalla hasta determinar que disc jockey divierte más a la multitud que lo escucha [3], [4]. En trabajos previos del ISIER-UM¹ [5], [6], [7], [8] orientados al control de Robots y el control de artefactos en el contexto de la domótica [9], como así también en la lectura emocional del usuario, enfocando la lectura de la excitación y meditación [10], se experimentó con BCI en particular con EMOTIV [11] integrando la respuesta de biopotenciales eléctricos de individuos para el control de robots a través del electro-miograma, electro-encefalograma y electro-oculograma que son bioseñales eléctricas generadas por los patrones de actividad de los músculos, el cerebro y los ojos del usuario. Los trabajos desarrollados en IFiNe-UM² en relación al registro de actividad bioeléctrica cerebral [12], [13] y de otros parámetros fisiológicos, tanto en humanos como en animales de experimentación, contribuyen directamente a la investigación y registros de parámetros biométricos. Las técnicas de inferencia del estado emocional de un ser humano a partir de biomarcadores sin que este deba controlar su propio estado eléctrico

cerebral representan un área relativamente novedosa de investigación que ha sido identificada como BCI pasivo [14]. Los sistemas BCI pasivos en el contexto de monitoreo de operadores en ambientes virtuales necesitan recabar información biometría, más allá de la potencialmente necesaria para inferir estados relacionados con emociones/sentimientos. Ejemplos son el reconocimiento de errores (potenciales de negatividad relacionada a errores o ERNs), vigilia relajada (ritmo alfa), compromiso con una tarea (calculado a partir de relaciones entre ritmos cerebrales [15]), carga de trabajo [16] o intención [17]. Los problemas que el diseño de sistemas BCI pasivos plantea no han sido resueltos aún de manera completamente satisfactoria [18]. Entre estos desafíos, encontramos: el procesamiento/filtrado de artefactos, el registro y análisis multi-escala, el problema inverso del EEG, el promediado de respuestas y el basamento neurocientífico, sin embargo su aplicación se extiende frente al empleo de sistemas invasivos que recurrieron a implantar electrodos intracraneales en la corteza motora de primates [19], [20].

Líneas de Investigación, Desarrollo e Innovación

En particular, nuestra investigación se concentra en el desarrollo de sistemas con el empleo de BCI pasivos que puedan evaluar las reacciones de una persona que está siendo entrenada en un ambiente virtual para ejecutar tareas críticas. La investigación aplicará una interfaz BCI

que permita convertir el EEG y el registro de otros biomarcadores, inducido por los cambios motivacionales y emocionales del usuario, en señales para ser analizadas y procesadas por un sistema computarizado. Los sistemas de simulación son entornos virtuales altamente realistas que por sus características brindan condiciones apropiadas para el estudio de las personas que interactúan con ellos. Para llevar adelante este proyecto se prevén líneas de desarrollo orientadas a: a) analizar dispositivos de bioseñales y sensores biométricos de parámetros fisiológicos; b) experimentar con interfase cerebro-máquina; c) desarrollar el Framework de monitoreo de bioseñales; e) efectuar la experimentación basal; f) realizar la integración con el ambiente de simulación contemplando captura de situaciones, parámetros de las mismas y acciones del usuario; g) generar la trazabilidad del estado emocional del individuo en función de la evolución de las situaciones de simulación; h) efectuar la experimentación integrada en contexto de simulación específicos y analizar sus resultados, i) generar un perfil del individuo en un contexto de computación afectiva.

Objetivos y Resultados Esperados

Con la ejecución del presente proyecto se pretende contribuir a inferir los estados anímicos de los usuarios, durante su actuación en los sistemas de adiestramiento y a otros entornos virtuales de aprendizajes, como así también asistir al estudio de los parámetros biométricos más relevantes.

La experimentación inicial abordará situaciones de simulación basal a situaciones de casos de borde donde se registre los cambios biométricos que permitirá deducir el estado emocional del individuo en situaciones esperadas en términos de relajación o en situaciones esperadas en términos de excitación, entre otras que caracterizan los cambios motivacionales y emocionales. La relevancia de este tipo de investigación también puede ser analizada bajo la perspectiva de una mejor comprensión del comportamiento de las personas en determinadas situaciones en donde se realizarán los ensayos. Esto permitirá analizar de modo crítico los modos de capacitación con sistemas de simulación mediante realidad virtual. La evaluación de las emociones humanas en un sistema de simulación permitirá obtener información más precisa que incida, también, en los modos de enseñar, capacitar y aprehender mediante sistemas de simulación. Con la ejecución este proyecto se pretende mejorar los procesos de capacitación y evaluación de sujetos expuestos a situaciones críticas. La propuesta es registrar indicadores biométricos durante las sesiones que desarrolla un usuario en ambientes virtuales representados por sistemas de simuladores y adiestradores, entre otros, con especial interés en sistemas que asisten al entrenamiento de personal en sistemas de misión crítica, como así también en sistema que por su naturaleza puedan influir en la seguridad social. Los actuales sistemas en uso para el adiestramiento en sistemas virtuales recrea las escenas y situaciones, pero no

registran información biométrica relacionada especialmente con el estado emocional, situación de excitación-relajación con el empleo de una interfaz Cerebro-Máquina (BCI) que permita convertir parámetros fisiológicos, cognitivos y emocionales del usuario (entre otras, la actividad eléctrica cerebral (EEG), parámetros de pulsaciones, presión arterial, frecuencia cardiaca) en especial frente a situaciones simuladas críticas como lo son fallos y su resolución, aplicación de procedimientos de emergencia, a fin de poder evaluar no solo la resolución del caso en forma aislada, sino también el comportamiento biométrico del individuo frente a la situación presentada en el contexto virtual, contribuyendo a lograr un perfil y su historial comparativo en sus distintas sesiones de entrenamiento. De tal forma, al recolectar y procesar información del estado del usuario (a partir de biomarcadores) se busca evaluar/mejorar momento a momento la interacción entre el hombre y el ambiente virtual. Se considera en este orden la influencia en los resultados de entrenamiento del individuo en función de los cambios emocionales y patrones biométricos en función del contexto de simulación. Los cambios motivacionales y emocionales, representar la alteración de señales bioeléctricas contrastándolas y determinando cambios emocionales en función del contexto de actuación del usuario. Se presenta el modelo de integración de información biométrica en función del mundo virtual (Fig. 1).

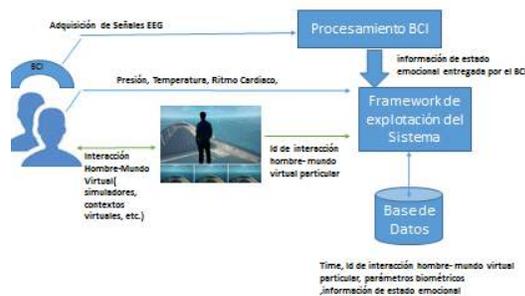


Fig 1. Modelo conceptual del sistema de influencia del estado biométrico emocional de personas interactuando en mundos virtuales

Formación de Recursos Humanos

El grupo de investigación se encuentra conformado por tres investigadores formados, dos investigadores en formación, se prevé la incorporación de dos estudiantes grado durante el año 2018.

Referencias

1. Hamadicharef, 2010 "Brain Computer Interface Literature- A bibliometric study", in 10th International Conference on Information Science, Signal Processing and their Applications, Kuala Lumpur, pp. 626-629.
2. J. del R Millán, "Brain-computer interfaces, 2002" in Handbook of Brain Theory and Neural Networks, 2nd ed, M.A. Arbib, Ed. Cambridge, MA: MIT Press.
3. M. A. Lebedev and M. A. L. Nicolelis, 2006. "Brainmachine interfaces: Past, present and future," Trends in Neurosciences, vol. 29, no. 9, pp. 536-546.
4. R. Wolpaw, D. J. McFarland, 2000 "Brain-computer interface research at the Wadsworth center," IEEE Trans. Rehab. Eng., vol. 8, pp. 222-226.
5. Ierache, J., Dittler M., Pereira G., García Martínez R, "Robot Control on the basis of Bio-electrical signals" XV CACIC 2009,

- UNJu, ISBN 978-897-24068-3-9
6. Ierache J, Pereira G, Iribarren J, Sattolo I, "Robot Control on the Basis of Bio-electrical Signals": "International Conference on Robot Intelligence Technology and Applications" (RiTA 2012). Series Advances in Intelligent and Soft Computing of Springer.
 7. Ierache., J, Pereira.,G, Sattolo.,I ,Guerrero., A, D'Alto J, Iribarren., J. Control vía Internet de un Robot ubicado en un sitio remoto aplicando una Interfase Cerebro-Máquina".XVII CACIC 2011, UNLP, ISBN 978-950-34-0756-1,p 1373-1382.
 8. Ierache J, Pereira G, Iribarren J "Demostración de los resultados en la integración de Interfases Lectoras de Bioseñales aplicadas al Control de un Robot" VII TEYET 2012 UNNOBA, 2012, demos educativas. ISBN 978-987-28186-3-0.
 9. Ierache., J, Pereira., G, J, Iribarren Framework for Brain Computer Interface implemented to control devices in the context of home automation XIX CACIC 2013, CAECE Mar del Plata, ISBN 978-897-23963-12
 10. Ierache J, Nervo F, Pereira G, Iribarren J Estado emocional centrado en estímulos, aplicando Interfase cerebro-maquina, XX Congreso Argentino de Ciencias de la Computación (Buenos Aires, 2014) ISBN 978-987-3806-05-6
 11. Emotiv Epoc Vigente 2018 <http://www.emotiv.com/>.
 12. Cervino, C. 1999. MS. Nuevo ritmo alfa y actividad gamma en los bulbos olfatorios de un armadillo sudamericano: cuantificación y perspectivas futuras. Premio "Prof. Dr. Eduardo de Robertis", al mejor trabajo de Neurociencia Básica. 455 pág.
 13. Affanni, JM. y CO. Cervino. 2000. Actividad Bioeléctrica Cerebral. En: Fisiología Humana de Houssay (7ma ed.), A. Houssay y C. Cingolani (eds.). Buenos Aires: Ed. El Ateneo. 1120 pág.: 1015-1039.
 14. T. Zander y C. Kothe. 2010. Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *J. Neural Eng.*, 8(2)-025005 <http://dx.doi.org/10.1088/1741-2560/8/2/025005>.
 15. L. Georgey A. Lécuyer. 2010. An overview of research on "passive" brain-computer interfaces for implicit human-computer interaction. International Conference on Applied Bionics and Biomechanics ICABB 2010 - Workshop W1 "Brain-Computer Interfacing and Virtual Reality", Venecia, Italia.
 16. D. Heger, F. Putze, F. y T. Schultz, T. 2010. Online Workload Recognition from EEG data during Cognitive Tests and Human-Computer Interaction. *KI 2010: Advances in Artificial Intelligence*, 410-417.
 17. S. Haufe, J.W. Kim, I.H. Kim, A. Sonnleitner, M. Schrauf, G. Curio, B. Blankertz. 2014. Electrophysiology-based detection of emergency braking intention in real-world driving. *J. Neural Eng.* 11(5)-056011. [10.1088/1741-2560/11/5/056011](https://doi.org/10.1088/1741-2560/11/5/056011)
 18. S. Makeig, C. Kothe, T. Mullen, N. Bigdely-Shamlo, Z. Zhang y K. Kreutz-Delgado. 2012. Evolving Signal Processing for Brain-Computer Interfaces. *Proceedings of the IEEE - Volume 100 (Special Centennial Issue)*.
 19. J. Wessberg, C. R. Stambaugh, J. D. Kralik, P. D. Beck, M. Laubach, J. K., , 2000. "Real-time prediction of hand trajectory by ensembles of cortical neurons in primates," *Nature*, vol. 408, pp. 361-365.
 20. M. A. L. Nicolelis, "Brain-machine interfaces to restore motor function and probe neural circuits," *Nature Rev.Neurosci.*, vol. 4, pp. 417-422, 2003

Aplicaciones Móviles 3D

Pablo Thomas, Federico Cristina, Sebastián Dapoto, Patricia Pesado

Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – Universidad Nacional de La Plata

50 y 120 La Plata Buenos Aires

Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

{pthomas, fcristina, sdapoto, ppesado}@lidi.info.unlp.edu.ar

Resumen

Se presenta una línea de investigación y desarrollo, que tiene por objeto estudiar temas relacionados con aspectos de Ingeniería de Software, orientados al desarrollo de aplicaciones móviles sobre diversas plataformas y entornos operativos. En particular, se pone el énfasis en el desarrollo de herramientas para dispositivos móviles que utilicen un entorno tridimensional (3D).

Palabras claves: Dispositivos Móviles - Aplicaciones 3D - Aplicaciones Multiplataforma - M-Learning – Performance – Consumo de energía

Contexto

Esta línea de Investigación forma parte del proyecto (2018-2021) “*Metodologías, técnicas y herramientas de ingeniería de software en escenarios híbridos. Mejora de proceso.*”, en particular del subproyecto “*Ingeniería de Software para escenarios híbridos*”, del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación de la Nación.

Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa en proyectos financiados por el Ministerio de Ciencia y Tecnología de España y la AECID.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Por otra parte, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado.

Introducción

Los dispositivos móviles son cada vez más sofisticados y su evolución tecnológica permite ejecutar aplicaciones cada vez más complejas y con exigentes requerimientos de hardware. Debido a esto, actualmente existe una gran cantidad de librerías y frameworks que permiten desarrollar aplicaciones tridimensionales para dispositivos móviles.

Las aplicaciones desarrolladas mediante estos frameworks son visualmente más agradables y generan una mejor experiencia de usuario, dado que los entornos tridimensionales son más cercanos a la realidad que los bidimensionales, y permiten involucrarse de forma más activa.

Por otro parte, el ámbito educativo debe adaptarse a los cambios y nuevas formas de aprendizaje. M-learning (mobile learning) plantea métodos modernos de apoyo al proceso de aprendizaje mediante el uso de dispositivos móviles. Las aplicaciones móviles 3D son una herramienta ideal para acercar a los alumnos.

Sin embargo, muchos de los potenciales usuarios de estas aplicaciones educativas pueden no disponer de dispositivos de última generación. Por esta razón, resulta de vital importancia realizar un análisis en profundidad de los parámetros que inciden en la performance final de una aplicación 3D.

Otro aspecto de relevancia es el consumo de energía de este tipo de aplicaciones, el cual suele ser bastante alto. Teniendo en cuenta esto, es importante realizar un estudio que permita identificar qué características de este tipo de aplicaciones producen mayor consumo de energía.

Líneas de Investigación y Desarrollo

- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para dispositivos móviles.
- Mobile Learning
- Aplicaciones Móviles 3D Multiplataforma
- Frameworks para el desarrollo de Aplicaciones Móviles 3D
- Performance de aplicaciones móviles 3D
- Consumo de energía en aplicaciones móviles 3D

Resultados esperados/obtenidos

Los resultados esperados/obtenidos se pueden resumir en:

- Avanzar en la capacitación continua de los miembros de la línea de investigación.
- Avanzar en el aprendizaje de frameworks que permiten desarrollar aplicaciones 3D multiplataforma, particularmente para dispositivos móviles.
- Avanzar en el desarrollo de aplicaciones educativas, teniendo como finalidad enriquecer las experiencias interactivas y motivar el aprendizaje mediante su uso.
- Avanzar en el análisis de performance de uno o más frameworks 3D con el fin de detectar los puntos críticos y mejorar la ejecución de las aplicaciones 3D.
- Avanzar en el análisis de consumo de energía de aplicaciones generadas con diferentes frameworks 3D.
- Se han evaluado diferentes frameworks para el desarrollo de aplicaciones móviles 3D.

- Se ha desarrollado el prototipo móvil R-Info3D, similar a la aplicación de escritorio R-Info, que mejora la experiencia del alumno y permite una mejor visualización de las tareas que realiza el robot desde diferentes puntos de vista.
- Se ha desarrollado el prototipo móvil InfoUNLP3D que sirve de guía para los estudiantes en sus primeras experiencias dentro de la Facultad de Informática. Figura 1.



Figura 1. InfoUNLP3D.

- Se ha desarrollado mediante Unity Engine un prototipo de análisis de performance de las aplicaciones 3D, que estudia la degradación de la ejecución de dichas aplicaciones a medida que se incrementa ciertos factores de relevancia como lo son la cantidad de objetos simples y complejos, objetos con y sin textura, luces y sombras, y sistemas de partículas. Figuras 2, 3 y 4.
- Se está avanzando en el desarrollo de un prototipo de análisis de performance de las aplicaciones 3D desarrolladas mediante el framework Unreal Engine y/o el framework CryEngine, con el fin de realizar un estudio comparativo con el prototipo desarrollado para Unity.

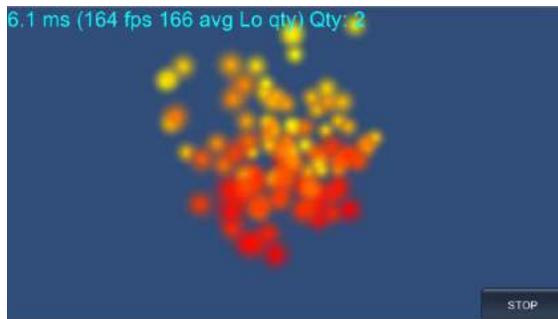


Figura 2. Prototipo de Performance. Prueba de sistema de partículas.

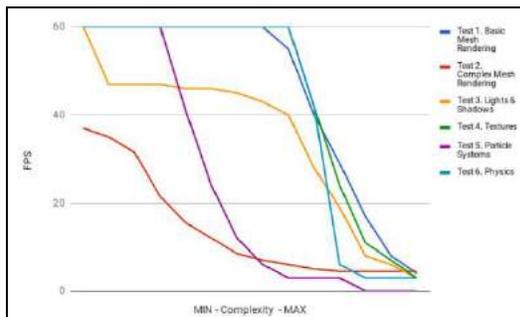


Figura 3. Evolución de los cuadros por segundo (FPS, sigla en inglés) en cada prueba.

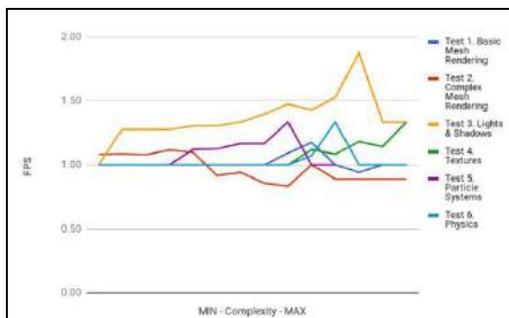


Figura 4. Comparación de resultados en diferentes calidades de renderizado.

Formación de Recursos Humanos

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con el proyecto. Además participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

Referencias

1. Federico Cristina, Sebastián Dapoto, Pablo Thomas, Patricia Pesado. Capítulo de Libro: "Performance evaluation of a 3D engine for mobile devices". Libro: "Computer Science – CACIC 2017. Communications in Computer and Information Science, vol 790". Springer. ISBN: 978-3-319-75213-6, 978-3-319-75214-3, páginas 155-163. Año 2018.
2. Federico Cristina, Sebastián Dapoto, Pablo Thomas, Patricia Pesado. Capítulo de Libro: "InfoUNLP3D: An interactive experience for freshman students". Libro: "Computer Science & Technology Series - XXII Argentine Congress of Computer Science. Selected Papers" (362 páginas). EDULP. ISBN: 978-987-4127-28-0, páginas 249-256. Año 2017.
3. Federico Cristina, Sebastián Dapoto, Pablo Thomas, Patricia Pesado. Capítulo de Libro: "3D Mobile Prototype for Basic Algorithms Learning". Libro: "Computer Science & Technology Series - XXI Argentine Congress Of Computer Science. Selected Papers" (300 páginas). EDULP. ISBN: 978-987-4127-00-6, páginas 239-247. Año 2016.
4. Linowes J. "Unity Virtual Reality Projects". 2015. ISBN-13: 978-1783988556.
5. Akekarat Pattrasitidecha. "Comparison and evaluation of 3D mobile game engines". Chalmers University of Technology. University of Gothenburg. 2014.
6. De Giusti Laura, Leibovich Fabiana, Sánchez Mariano, Rodríguez Eguren Sebastian, Chichizola Franco, Naiouf Marcelo, De Giusti Armando. "Herramienta interactiva para la enseñanza temprana de Concurrencia y Paralelismo: un caso de estudio". CACIC 2014. San Justo. ISBN: 978-987-3806-05-6.
7. Aaron Carroll, Gernot Heiser. "An Analysis of Power Consumption in a Smartphone". USENIXATC'10 Proceedings of the 2010 USENIX conference on USENIX annual technical conference. 2010.
8. Kantel E., Tovar G., Serrano A. "Diseño de un Entorno Colaborativo Móvil para Apoyo al Aprendizaje a través de Dispositivos Móviles de Tercera Generación." IEEE-RITA 5, no. 4 (2010): 146-151.

9. Rosa Paredes, J. Alfredo Sánchez, Liliana Rojas, Daniel Strazzulla, Ronel Martínez-Teutle. "Interacting with 3D Learning Objects". 2009 LA Web Congress. ISBN: 978-0-7695-3856-3/09.
10. Unity 3D Homepage: <https://unity3d.com/>.
11. Optimizing Graphics Performance. <https://docs.unity3d.com/Manual/OptimizingGraphicsPerformance.html>
12. Unreal Engine Homepage: <https://www.unrealengine.com/>.
13. CryEngine. Homepage: <https://www.cryengine.com/>

Evaluación del desarrollo tecnológico para la definición de Industrias 4.0

Alicia Mon; Horacio Del Giorgio; Eduardo De María; Claudio Figuerola; Matías Querel

Departamento de Ingeniería e Investigaciones Tecnológicas

Universidad Nacional de La Matanza.

Florencio Varela 1903 - San Justo (CP 1754)

Tel: 4480-8952

alicialmon@gmail.com; hdelgiorgio@unlam.edu.ar; demaria.edu@gmail.com; claudio.figueroa@gmail.com; matias.querel@gmail.com

Resumen

El presente artículo expone los lineamientos de un proyecto cuyo objetivo reside en investigar sobre los tipos de TICs que implementan las industrias en sus diferentes procesos y áreas de negocios. Para ello se aplicará un conjunto de instrumentos diseñados por el Grupo GIS en un proyecto precedente, a efectos de realizar la evaluación en las industrias radicadas en el Partido de La Matanza. Como herramienta de soporte, se desarrollará una aplicación de software que permita relevar, analizar y evaluar la inserción de TICs de manera permanente.

A partir de dicha evaluación, podrán definirse los niveles del desarrollo tecnológico en cuanto a los productos software, hardware e infraestructura implementados por rama de actividad para poder establecer los parámetros y productos que definen a la Industria 4.0.

De este modo, se propone detectar los tipos de tecnologías instaladas en la actualidad, analizar el valor agregado de su utilización en los sectores industriales y determinar las necesidades de implementación de TICs en las cadenas de valor para confluir en la Industria 4.0.

Una vez determinados con precisión los parámetros tecnológicos de estas industrias, se trabajará sobre atributos de usabilidad que requieren las TICs para poder desarrollar productos centrados en los usuarios específicos.

Palabras clave: TICs, Industria 4.0, Ingeniería de Software, Usabilidad.

Contexto

En el año 2017 en la Universidad Nacional de La Matanza se inauguró el polo de desarrollo de software dependiente del Depto. de Ingeniería e Investigaciones Tecnológicas (DIIT) para la promoción y radicación de empresas de la industria de software, de modo tal de generar inserción laboral de los estudiantes de Ingeniería Informática. Es por ello que la investigación que se está desarrollando se propone como objetivo la generación de información del desarrollo de TICs (Tecnologías de la Información y las Comunicaciones) en la industria local que permita vincular a las empresas de software radicadas en la UNLaM con dicho sector y facilitar la detección de necesidades de desarrollo e implementación de productos software.

Esta línea, incluye un proyecto de investigación del DIIT dentro del programa de incentivos, un Proyecto PICTO financiado por el Ministerio de Ciencia y Tecnología y la vinculación con grupos de investigación de la región.

Sobre esta línea, el DIIT se propone estudiar los tipos de tecnologías instaladas en la actualidad, analizar el valor agregado del uso de estas tecnologías en los sectores industriales y determinar las necesidades de implementación de TICs en las cadenas de valor para confluir en la Industria 4.0., además de elaborar información pertinente para el desarrollo tecnológico, la inserción laboral de alumnos y graduados y la formación/actualización de los docentes de la Universidad.

Introducción

La escalada en la evolución de las TICs, en forma confluyente con el desarrollo de las distintas disciplinas científicas, impactan de lleno en las transformaciones de todos los aspectos que hacen a la producción de bienes y servicios, de modo tal que se está generando un proceso innovador que se lo reconoce como la cuarta revolución industrial [Ministerio de Ciencia, Tecnología e Innovación Productiva, 2015] [Ibáñez Díaz, Cabas Alonso, Cuevas Arce, & Balaguer, 2018] [Kantar Millward Brown, 2017]. Una serie de avances en diversos campos como la robótica, la biotecnología, la genética, la nanotecnología, la expansión de internet a internet de las cosas, el desarrollo de la inteligencia artificial y la fabricación aditiva (impresiones 3D) se suman a la revolución energética de fuentes renovables y el desarrollo de las TICs que caracterizaron a la tercera revolución industrial.

Este conjunto de descubrimientos científicos y tecnológicos fueron tomando la forma de innovación y comenzaron a manifestarse, no sólo en la vida cotidiana de las personas sino, en los procesos que se desarrollan en la actividad económica, tanto en la producción industrial de bienes, así como en la prestación de servicios. Dentro de esos grandes campos la digitalización ha cobrado un rol protagónico y en algunas áreas se ha tornado imprescindible. En este sentido, cobra fuerza el término Industria 4.0, que refiere específicamente a la cuarta revolución industrial e implica un salto cualitativo significativo en la organización y gestión de las cadenas de valor.

El desarrollo de las TICs permite la hibridación entre el mundo físico y el digital; es decir que posibilitan la vinculación del mundo físico a través de dispositivos, materiales, productos, maquinaria e instalaciones, con el mundo digital, representado por sistemas y productos software. Esta conexión habilita que los dispositivos y sistemas colaboren

entre ellos y con otros sistemas para crear una industria inteligente.

La Industria 4.0 representa la integración de extremo a extremo de la cadena de valor que va desde los cambios de demandas del gran público al logro de su satisfacción por parte de las fábricas inteligentes. Ya no tendrá sentido hablar de simples fábricas. Las fábricas serán inteligentes (Smart factories) y llegará el día en que no tendrá sostenibilidad una fábrica que no se haya adaptado a la cuarta generación.

Frente a esta gran transformación, la Industria actual necesita cambios tecnológicos urgentes, dado que la competitividad de las empresas pasa por la globalización, la productividad y la innovación. Sin embargo, no se han encontrado trabajos que definan con exactitud cuál es el grado de desarrollo tecnológico que se encuentra hoy en día implementado en la industria para poder determinar cuáles son los requerimientos específicos de actualización.

En este contexto no se ha encontrado, en la bibliografía actual, las características específicas de productos software, hardware y comunicaciones que definan con precisión los atributos de las Industrias 4.0. Si bien, las nuevas herramientas, las nuevas tecnologías, los nuevos materiales, las nuevas metodologías, las nuevas fuentes de energía y todos los factores que se engloban bajo el nombre de Industria 4.0 constituyen las palancas imprescindibles para lograrlo, no se ha encontrado de manera precisa la definición de dichos elementos.

Por otra parte, la Ingeniería de software se encuentra trabajando fuertemente en la interacción entre el humano y la computadora, área de trabajo conocida como HCI (Human Computer Interaction), que permite abordar el diseño de productos software centrados en los usuarios específicos de cada producto en su contexto real de uso.

Si bien los sistemas software están dirigidos a un público cada vez más amplio,

a usuarios cada vez menos expertos en el manejo de sistemas informáticos, los softwares que se utilizan en la industria son complejos y en la medida que su interfaz con el usuario resulte tan compleja como el producto, su usabilidad será deficiente y la potencialidad del producto quedará restringida a la capacidad y el expertise de los usuarios, que muy habitualmente no se los califica para ese tipo de desarrollos tecnológicos. Esto implica que el diseño de productos software con alta usabilidad constituye un atributo fundamental para el éxito de un producto TIC.

Líneas de Investigación, Desarrollo e Innovación

La posibilidad de conocer las diferentes tecnologías, los tipos de productos software instalados, así como la agregación de valor que aportan en la productividad, resulta una información clave para la toma de decisiones estratégicas tanto en la industria del software como en los diferentes sectores industriales.

El presente proyecto, que se ha iniciado en el Departamento de Ingeniería e Investigaciones Tecnológicas de la UNLaM (DIIT), se enfoca en las líneas de investigación que el grupo GIS viene desarrollando desde hace varios años y en red con otras universidades en el marco de la ingeniería de software, el desarrollo de TICs en la industria y los desarrollos hacia la industria 4.0.

Dichos proyectos se han propuesto investigar un método de cálculo para un índice de implantación de TICs (Software, Hardware e Infraestructura) que se utilizan en las diferentes ramas de actividad industrial.

Los proyectos precedentes han generado variados resultados científicos y académicos publicados en diferentes congresos nacionales e internacionales y la interdisciplinariedad del mismo ha facilitado el desarrollo de una Tesis del

Doctorado en Ciencias Económicas de la UNLaM que ha permitido generar información relevante sobre el desarrollo local para la investigación. La tesis *Exploración de la inserción de las Tecnologías de la Información y Comunicación en el desarrollo industrial*, ha sido escrita por el Mg. Horacio René Del Giorgio y dirigida por la Dra. Alicia Mon, en la Escuela de Posgrado que se encuentra en proceso de evaluación final para su defensa en el transcurso del año 2018.

Para el desarrollo del presente proyecto, el grupo GIS propone conformarse como un grupo interdisciplinario, integrado por Ingenieros Informáticos, Industriales y Electrónicos, que aportarán un amplio conocimiento sobre los procesos industriales, sobre las tecnologías aplicadas en diversos sectores productivos y sobre las nuevas tendencias en desarrollos tecnológicos para la confluencia hacia la Industria 4.0. Dichos profesionales conforman actualmente la planta de docentes-investigadores de la UNLaM.

Asimismo, en los proyectos de investigación antes mencionados, se ha estudiado la conformación industrial del Partido de La Matanza, siendo la vinculación con el desarrollo local un factor predominante en las líneas de investigación desarrolladas desde el Departamento.

Resultados y Objetivos

El desarrollo de las TICs encamina a su vez el desarrollo de la industria hacia la convergencia digital, la conectividad entre objetos y la incorporación de inteligencia artificial en la resolución de problemas de producción, entre otros aspectos que enfocan hacia lo que se conoce como Industria 4.0.

Este tipo de industrias encierran un conjunto de desarrollos tecnológicos sin definiciones precisas, y es por ello que la problemática a investigar en este proyecto se refiere a determinar específicamente

cuáles son los avances tecnológicos que específicamente definen a una industria como 4.0. El grupo de investigación GIS ha desarrollado un método de medición de inserción de TICs en la industria que permite determinar los tipos de productos software, hardware y comunicaciones que pueden ser incorporados por una industria según su área funcional. Este método permite evaluar en 3 niveles diferenciados de desarrollo tecnológico según las TICs como básico, medio o avanzado. En el tipo de TICs que se agrupan en el nivel avanzado se encontrarían las industrias más desarrolladas tecnológicamente, sin que necesariamente lleguen a ser reconocidas como Industrias 4.0.

Es por ello que la investigación se propone determinar con precisión los tipos de TICs que debe contener una industria para poder definirla como una industria 4.0, y asimismo se propone establecer cuáles son los criterios de usabilidad que deberían cumplir dichas tecnologías para ser sencillas de utilizar por los usuarios específicos en su contexto real de uso al interior de las industrias.

El proyecto se propone desarrollar una aplicación de software que permita realizar un relevamiento y ordenamiento de la información sobre las TICs en la industria del Partido de La Matanza, a efectos de realizar un análisis, evaluación y medición de TICs para definir los niveles de desarrollo tecnológicos en la actualidad y las características específicas de las industrias 4.0, así como los atributos de usabilidad requeridos.

A su vez, en la Universidad Nacional de La Matanza se ha creado un Polo de desarrollo tecnológico en el que se promueve la radicación de empresas de la industria de software a efectos de generar una inserción laboral local de los estudiantes y graduados de las carreras de Ingeniería. La definición de instrumentos de relevamiento sistemático y permanente se propone como inicio para conocer las industrias de la zona, indagar sobre las

tecnologías que tienen incorporadas en diferentes áreas del proceso productivo, conocer y analizar las causas de los límites para incorporar mayor tecnología y construir información que permita detectar necesidades de nuevos desarrollos para colaborar en el desarrollo de las industrias 4.0.

Formación de Recursos Humanos

El grupo de investigación GIS se ha conformado para este proyecto como un grupo interdisciplinario e interuniversitario, integrado por Ingenieros Informáticos, Industriales y Electrónicos, todos docentes-investigadores.

Un integrante del grupo GIS se encuentra desarrollando una tesis de la Maestría en Dirección Estratégica y Tecnológica del Instituto Tecnológico Buenos Aires, en tanto que otro de los investigadores se encuentra desarrollando su tesis del Doctorado en Ciencias Económicas de la UNLaM sobre el tema abordado en este Proyecto.

Se prevé la incorporación de 4 alumnos de grado para que realicen su proyecto final de carrera en el marco del proyecto.

Bibliografía

- ANETCOM. (2017). La TIC en la estrategia empresarial. Valencia.
- Federal Ministry of Education and Research. (2018). High Tech and Innovation. Obtenido de <https://www.bmbf.de/en/the-new-high-tech-strategy-2322.html>
- Ferré Grau, X. (2014). Incrementos de Usabilidad al Proceso de Desarrollo Software. Obtenido de <https://www.researchgate.net/publication/221595137>
- Ferré Grau, X. (2014). Principios Básicos de Usabilidad para Ingenieros Software. Obtenido de <https://www.researchgate.net/publication>

[n/221595210 Principios Basicos de Usabilidad para Ingenieros Software](#)

- Ibáñez Díaz, A., Cabas Alonso, J., Cuevas Arce, S., & Balaguer, C. (2018). Observatorio de la Industria 4.0 - Foro de Profesionales. Obtenido de <http://www.observatorioindustria.org/Inicio/>
- Kantar Millward Brown. (2017). Tercer Estudio de Competencias Digitales en la Empresa Española. Obtenido de <http://www.ticpymes.es/siteresources/files/839/54.pdf>
- Ministerio de Ciencia, T. e. (2015). Industria 4.0: Escenarios e impactos para la formulación de políticas tecnológicas en los umbrales de la Cuarta Revolución Industrial. Obtenido de <http://www.mincyt.gob.ar/adjuntos/archivos/000/038/0000038319.pdf>
- Serra, D., Rodríguez, S., Novellino, H., Boychenko, D., Penella, C., & Incaugarat, N. (2016). Caracterización del perfil exportador de las pymes industriales del partido de La Matanza. Lomas de Zamora - Provincia de Buenos Aires: Instituto de Investigaciones en Ingeniería Industrial - Universidad Nacional de Lomas de Zamora.

Hacia la Formalización de un Lenguaje Visual Unificador de UML, EER y ORM 2

Federico Solorza¹

Germán Braun^{1,2,3}
Pablo Fillottrani^{2,4}

Laura Cecchi¹

email: federico.solorza@est.fi.uncoma.edu.ar,
{german.braun,lcecchi}@fi.uncoma.edu.ar, prf@cs.uns.edu.ar

¹*Grupo de Investigación en Lenguajes e Inteligencia Artificial*
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

²*Laboratorio de I&D en Ingeniería de Software y Sistemas de Información*
Departamento de Ciencias e Ingeniería de la Computación
UNIVERSIDAD NACIONAL DEL SUR

³*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*

⁴*Comisión de Investigaciones Científicas de la provincia de Buenos Aires (CIC)*

Resumen

La presente investigación se desarrolla mediante el trabajo de docentes investigadores de la Universidad Nacional del Comahue (UNCo) y de la Universidad Nacional del Sur (UNS), en el contexto de proyectos de investigación financiados por las universidades indicadas.

El objetivo general del presente trabajo consiste en definir, diseñar, especificar y formalizar un lenguaje de modelado gráfico que permita unificar el uso de los lenguajes de modelado UML, ER, EER y ORM 2. La idea es proporcionar un mecanismo de abstracción respecto de un metamodelo integrador, sobre el cual nuestro lenguaje estará basado, facilitando el diseño de modelos conceptuales y la edición y visualización de ontologías en el contexto una herramienta gráfica Web con asistencia de razonamiento automática.

Palabras Clave: Ingeniería de Software basada en Conocimiento, Lógicas Descriptivas, Ontologías, Interoperabilidad de Lenguajes de Modelado Conceptual.

Contexto

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el marco del proyecto de investigación *Agentes Inteligentes y Web Semántica (04/F014)*, por la Universidad Nacional del Sur a través del proyecto de investigación *Integración de Información y Servicios en la Web (24/N027)* y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral. Los proyectos de investigación tienen una duración de cuatro años y la beca doctoral una duración de 5 años, finalizando esta última en abril de 2019.

1. Introducción

Compartir ontologías y determinar la consistencia de su modelo a nivel de la capa conceptual no es un trabajo trivial, más aún cuando dicho modelo deviene de la integración de otros modelos, posiblemente diseñados utilizando otros lenguajes conceptuales o incluso

cuando la realización de los modelos proviene de fuentes heterogéneas. Este tipo de situaciones ocurre en la integración de modelos ontológicos desarrollados por especialistas de diversas ciencias (biología [1], medicina, física, química), iniciativas e-government e incluso cuando dos compañías se fusionan [2].

Para poder superar esta complejidad, se requiere de un metamodelo que unifique a aquellos diseñados en diferentes lenguajes. A partir de la evaluación de los lenguajes de modelado conceptual (de ahora en más CMD) más utilizados (UML, ER, EER, ORM2) y del análisis de sus características, constructores y demás elementos con un enfoque que permita entender la concepción ontológica que les dió origen, surge el metamodelo KF [3]. El metamodelo KF unifica los aspectos estáticos de los lenguajes UML v2.4.1, ER, EER, ORM y ORM2, de tal manera que subsume a cada uno de ellos en algún fragmento del metamodelo.

Tanto en el diseño de modelos conceptuales tales como Bases de Datos, como en el diseño ontológico, los lenguajes gráficos ofrecen abstracciones que permiten sobrellevar la complejidad inherente a estos modelos. La habilidad de una notación gráfica para habilitar la detección temprana de posibles errores es clave en la construcción de un sistema complejo donde intervienen varias personas con diferentes perfiles de conocimientos.

El modelado conceptual es una tarea cognitiva relacionada al proceso de comprender y aprender sobre modelos reales [4]. En este contexto, la visualización tiene un rol crucial ayudando a la percepción de propiedades y patrones que no han sido previamente capturados. Además, permite identificar problemas sobre los datos y focalizar la atención sobre el dominio que está siendo modelado.

Bajo esta línea de investigación se propone definir, especificar y formalizar un lenguaje gráfico para el metamodelo KF. Hasta donde conocemos ningún lenguaje visual de estas características ha sido especificado. Asimismo, se implementará este nuevo lenguaje en la herramienta Web de diseño ontológico *crowd* [5, 6], para la cual un proceso de visualización que integra lenguajes gráficos con razonamiento ba-

sado en lógica, ha sido formalizado por autores del presente trabajo [7]. *crowd*, es una herramienta web cliente-servidor creada para sobrellevar la complejidad inherente que supone el diseño de un modelo conceptual y ontológico. Entre otros lenguajes y tecnologías, *crowd* soporta la integración de múltiples razonadores lógicos [8] mediante el protocolo OWLlink [9]. En los trabajos [5, 6], se presentó la arquitectura Web de *crowd* y un prototipo permitiendo, en primer instancia, asegurar que las diversas restricciones entre las entidades de un representante de una ontología sean lógicas e incluso materializables. El *front-end* de la herramienta permite al usuario escribir su modelo gráficamente, usando diagramas de clases UML mientras que su *back-end* se ocupa de validar los modelos e inferir restricciones que puedan estar implícitas.

Este trabajo consta de cuatro partes siendo esta introducción la primera de ellas. En la siguiente sección, se presentan los objetivos de los proyectos de investigación en los que se enmarca este trabajo y se describe la línea de investigación actual. En la sección 3 se indican algunos resultados obtenidos y trabajos futuros. Finalmente, en la última sección se comentan aspectos referentes a la formación de recursos humanos en esta temática.

2. Línea de Investigación y Desarrollo

La formación y puesta en valor de los conocimiento adquiridos y generados por medio del estudio especializado se enmarca en el proyecto de investigación Agentes Inteligentes y Web Semántica, de la UNCo, específicamente en el área de agentes. En este contexto, la representación y uso del conocimiento, por parte de dichos agentes en sistemas basados en la web, también nombrados bajo el título de Web Semántica son claves para la investigación. Esto involucra el aprendizaje y estudio acerca de las técnicas de representación de conocimiento y razonamiento, metodologías de modelado conceptual, estándares actuales referentes a la creación de lenguajes visuales, mecanismos para la in-

teroperabilidad de aplicaciones tanto a nivel de proceso como de datos. El empleo efectivo de los conceptos y conocimientos adquiridos dará soporte a comunidades de desarrollo de ontologías.

Por otro lado, en el proyecto de investigación Integración de Información y Servicios en la Web, UNS, se desarrollan metodologías y herramientas que asisten la interoperabilidad semántica de información y de servicios en la Web, privilegiando los últimos avances en el área de lenguajes de representación del conocimiento, ontologías y modelado conceptual.

Ambos proyectos comparten el perfil de investigación de este trabajo, en el que se estudian entre otros, sobre temas afines a la representación del conocimiento, las Lógicas Descriptivas [10], las Ontologías y aspectos semióticos [11], la Ingeniería de Software basada en Conocimiento y estándares para definir nuevos lenguajes visuales para representar conocimiento.

La OMG (Object Management Group) define varios lenguajes de modelado, dentro de los cuales UML [12] es muy utilizado en la actualidad. Existen otros de amplio uso, como EER [13] para bases de datos y ORM [14] para aquellos expertos en dominios e interacción entre el análisis de requerimientos y datos. Por esta razón, la interoperabilidad de los lenguajes de modelado conceptual se ha vuelto una necesidad, dado que los modeladores podrían requerir vincular entidades entre modelos representados en diferentes lenguajes de modelado conceptual y comunicarse con otros *skateholders* mediante lenguajes comunes.

OMG establece básicamente dos enfoques para la definición de lenguajes visuales. Una de ellas consiste en definir un nuevo lenguaje utilizando otro lenguaje especialmente diseñado para crear lenguajes de modelado conceptual basado en objetos. MOF (Meta Object Facility) [15] es un ejemplo de lenguaje creado de esa forma. La otra vía, consiste en modificar otro lenguaje de modelado (UML por lo general) utilizando mecanismos pre-establecidos. Agregados sintácticos y semánticos son posibles siempre que no se contradigan con el lenguaje modelo subyacente. Por ejemplo, las líneas definen relaciones en el lenguaje base, tal vez UML, y

por lo tanto no pueden denotar un aspecto totalmente diferente en el nuevo.

En esta línea de investigación se propone, como principal objetivo, definir, diseñar, especificar y formalizar un lenguaje de modelado gráfico tal, que permita unificar el uso de los lenguajes de modelado UML, ER, EER y ORM 2. La idea es que el nuevo lenguaje permita el uso mixto de los lenguajes de modelado facilitando el diseño de modelos conceptuales, la edición y visualización de ontologías y manteniendo, además, las funcionalidades de *crowd* basadas en la asistencia de razonamiento automático.

Existen herramientas que permiten el modelado conceptual para cada uno de los lenguajes de modelado mencionados anteriormente. Sin embargo, no hay herramientas que implementen al metamodelo tal como se propone en este trabajo.

3. Resultados Obtenidos y Trabajo Futuro

Inicialmente, para el desarrollo de un lenguaje gráfico para *crowd* se estudió y analizó el metamodelo KF que unifica a los tres lenguajes de modelado más usados tanto en su aspecto formal, expresado en las diversas lógicas descriptivas, como así también en lo ontológico e incluso relevancia cultural.

Diversos aspectos se contraponen entre los lenguajes UML, EER y se profundizan en ORM2 por ser un lenguaje sin atributos. Una de las metas para el diseño del lenguaje fue la importancia en los símbolos gráficos para denotar aspectos semióticos. ORM2 es un gran lenguaje de modelado que permite la verbalización de prácticamente todos sus constructores y posee además la variedad justa para la eficaz representación de intrincadas restricciones. UML y EER difieren de ORM2, ya que el mundo se representa principalmente por medio de entidades que cobran sentido en virtud de los atributos que se les especifica, y en EER un determinado conjunto de atributos, llamados colectivamente clave candidata, les proporciona identidad cuando todos ellos toman un valor en concreto. La Fi-

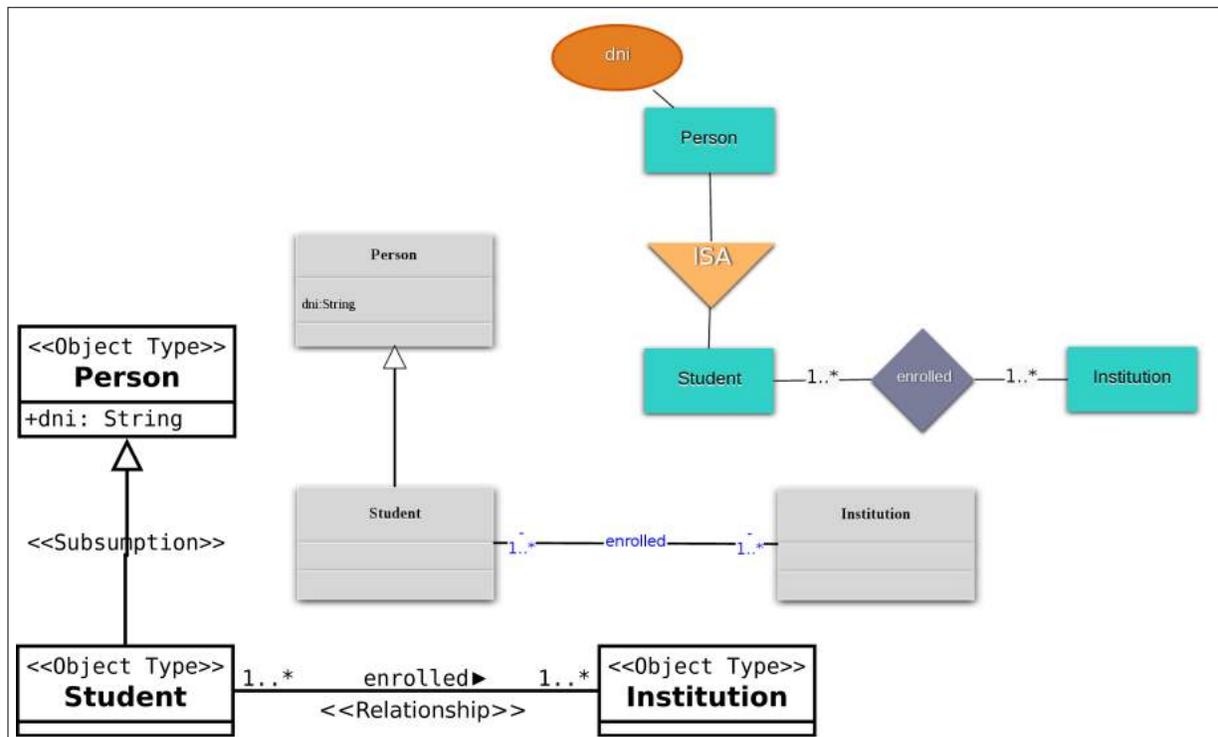


Figura 1: Enfoque preliminar de una sintaxis gráfica unificadora basada en perfiles UML.

Figura 1 muestra un enfoque preliminar basado en este análisis.

Hasta la fecha no se han desarrollado lenguajes gráficos que intenten unificar los lenguajes mencionados y estén basados en un metamodelo unificador. Existen, sin embargo trabajos que relacionan por ejemplo, ER y ORM [16].

4. Formación de Recursos Humanos

Sobre la temática de esta línea de investigación, uno de los autores de este trabajo está desarrollando su tesis de grado de la Licenciatura en Ciencias de la Computación.

Por otra parte, otro de los autores de este trabajo es becario doctoral CONICET y está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur.

Referencias

- [1] Gabriela Guardia, Ricardo Vêncio, and Clever De Farias. A uml profile for the obo relation ontology. *BMC Genomics*, 13:S3, Oct. 2012.
- [2] Albert Banal-Estañol. Information-sharing implications of horizontal mergers. *International Journal of Industrial Organization*, 25(1):31 – 49, 2007.
- [3] C. Maria Keet and Pablo Rubén Fillottrani. An ontology-driven unifying metamodel of uml class diagrams, eer, and orm2. *Data & Knowledge Engineering*, 98:30 – 53, 2015. Research on conceptual modeling.
- [4] Colin Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.
- [5] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Fillottrani. *Una Ar-* RedUNCI - UNNE - ISBN 978-987-3619-27-4

- quitectura Cliente-Servidor para Modelado Conceptual Asistido por Razonamiento Automático. In *XVIII Workshop de Investigadores en Ciencias de la Computación*, 2016.
- [6] Christian Gimenez, Germán Braun, Laura Cecchi, and Laura Fillottrani. crowd: A Tool for Conceptual Modelling assisted by Automated Reasoning - Preliminary Report. In *the 2nd Simposio Argentino de Ontologías y sus Aplicaciones SAOA '16 JAIIO '16*, 2016.
- [7] Germán Braun, Christian Gimenez, Laura Cecchi, and Pablo Fillottrani. Towards a Visualisation Process for Ontology-Based Conceptual Modelling. In *ONTOBRAS*, 2016.
- [8] V. Haarslev and R. Möller. Racer system description. In R. Goré, A. Leitsch, and T. Nipkow, editors, *International Joint Conference on Automated Reasoning, IJ-CAR'2001, June 18-23, Siena, Italy*, pages 701–705. Springer-Verlag, 2001.
- [9] Thorsten Liebig, Marko Luther, Olaf Noppens, and Michael Wessel. Owllink. *Semantic Web*, 2(1):23–32, 2011.
- [10] Diego Calvanese, Maurizio Lenzerini, and Daniele Nardi. Description logics for conceptual data modeling. In *Logics for Databases and Information Systems*, pages 229–263. Kluwer, 1998.
- [11] John F. Sowa. *The Role of Logic and Ontology in Language and Reasoning*, pages 231–263. Springer Netherlands, Dordrecht, 2010.
- [12] Grady Booch, James Rumbaugh, and Ivar Jacobson. *Unified Modeling Language User Guide*. Addison-Wesley Professional, 2005.
- [13] Martin Gogolla. *Extended Entity-Relationship Model: Fundamentals and Pragmatics*. Springer-Verlag, 1994.
- [14] Terry Halpin and Tony Morgan. *Information Modeling and Relational Databases*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2 edition, 2008.
- [15] OMG. MOF: Meta Object Facility, 2004. <http://www.omg.org/spec/MOF>, accedido en marzo de 2018.
- [16] John R. Venable and John C. Grundy. Integrating and supporting entity relationship and object role models. In Michael P. Papazoglou, editor, *OOER '95: Object-Oriented and Entity-Relationship Modeling*, pages 318–328, Berlin, Heidelberg, 1995. Springer Berlin Heidelberg.

JUEGO DE REALIDAD AUMENTADA PARA INCENTIVAR LA ACTIVIDAD FÍSICA EN NIÑOS CON DISCAPACIDAD

Cruz Alejandro^{1,3} & Acosta Nelson^{2,4}

¹ Departamento de Básicas y Aplicadas, Universidad Nacional de Chilecito, Chilecito, La Rioja

² Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires, Tandil, Buenos Aires

³ Becario Doctoral Agencia

⁴ Departamento de Ingeniería, Universidad Nacional de Tres de Febrero, Caseros, Buenos Aires

acruz@undec.edu.ar, el.nelson.acosta@gmail.com

RESUMEN

En la actualidad, con el avance de la tecnología, cada vez son más los niños que dedican su tiempo libre a actividades sedentarias. Esta problemática se acentúa aún más en los niños con diagnósticos dentro del espectro autista, diabetes, asperger, sobrepeso, obesidad infantil, entre otras, donde el sedentarismo tiene graves consecuencias. La actividad física durante los primeros años de vida es fundamental para el desarrollo físico, intelectual y social del niño. En este proyecto proponemos el desarrollo de un juego de realidad aumentada, compatible con el sistema operativo Android, para incentivar el hábito de la caminata en los niños jugando a la búsqueda del tesoro.

Palabras Clave: Serious gaming, Augmented Reality Games, Disability games.

CONTEXTO

El presente trabajo surge como una iniciativa para el desarrollo de herramientas que permitan mejorar la calidad de vida de los niños con discapacidad. Esta línea de trabajo se lleva a cabo en el ámbito de la empresa Tecnología LINDA S.R.L en conjunto con la Facultad de Cs. Exactas de la UNICEN.

1. INTRODUCCIÓN

La inactividad física se ha convertido en uno de los problemas de salud más importante tanto en los países desarrollados como en los países en desarrollo. La naturaleza de las actividades de ocio de los niños ha cambiado drásticamente en las últimas décadas. En el pasado, la infancia dedicaba gran parte de su tiempo de ocio a practicar juegos activos al aire libre. En la actualidad la evolución de la tecnología y la aparición de la televisión, consolas de juegos, pc, smartphones e internet, ha provocado que los niños dediquen una parte mucho mayor de su tiempo libre a actividades de tipo sedentarias. La inactividad física durante los primeros años de vida es un factor que colabora en el incremento de los niveles de obesidad y de otros trastornos médicos que se observan en niños, niñas y adolescentes [1, 2].

La actividad física influye sobre el funcionamiento psicológico de los individuos [3]. El deporte y el ejercicio proporcionan un medio importante para que los niños mejoren su bienestar social, su autoestima y sus percepciones sobre su imagen corporal, y su nivel de competencia, además de presentar más probabilidades de tener un mejor funcionamiento cognitivo. El ejercicio activo y habitual influye positivamente en la calidad de vida de las personas que padecen diversas

condiciones de salud. La mayoría de las personas con discapacidades o características especiales pueden beneficiarse de la práctica de ejercicio físico, adaptando la misma a los diferentes diagnósticos de cada individuo [4].

Esta investigación pretende “reintroducir” la actividad física en nuestros niños, utilizando como herramienta para incentivar el hábito de caminata, la actividad lúdica. El juego es una actividad recreativa que proporciona entretenimiento y diversión. El objetivo del juego no puede ser otra cosa que jugar. Lo más interesante del juego es que permite que la persona interactúe, asimile y comprenda las posibilidades y los límites que le ofrece determinado material o propuesta o en definitiva la realidad. El juego es una actividad exploradora, de aventura y experiencia, indispensable para el desarrollo físico, intelectual y social del niño [5].

Un enfoque educativo basado en los juegos electrónicos puede incluir características de entretenimiento y contenidos específicos para promover el proceso de aprendizaje del niño [6].

Tejeiro y Peregrina (2008)[7], proponen una definición para el término videojuego “entendemos por videojuego todo juego electrónico con objetivos esencialmente lúdicos, que se sirve de la tecnología informática y permite la interacción a tiempo real del jugador con la máquina, y en el que la acción se desarrolla fundamentalmente sobre un soporte visual (que puede ser la pantalla de una consola, de un ordenador personal, de un televisor o cualquier otro soporte semejante)”.

La industria de los videojuegos es una industria dedicada al ocio. Sin embargo en los últimos años ha tomado una importante relevancia investigaciones relacionadas a la creación de videojuegos orientados a la salud. Si bien estos juegos, en la mayoría de los casos, no tienen el presupuesto y los recursos comparados con los juegos comerciales que

encontramos en el mercado, logran aportar al jugador además de diversión y entretenimiento, algún tipo de beneficio o información sobre la salud de las personas. La mayoría de estos videojuegos acaban persiguiendo un fin educativo. Un serious games es un juego en el que la educación, en sus diversas formas, es el principal objetivo [8]. Estos juegos fomentan el aprendizaje y los cambios conductuales [6]. De una u otra manera, cuando se diseñan este tipo de videojuegos, se persigue la educación del jugador, tanto en la adquisición de hábitos saludables, como en la comprensión y entendimiento de procesos patológicos complejos. Investigadores en el campo de la salud han comenzado a aprovechar el poder de los videojuegos para motivar a los pacientes y, en última instancia, para mejorar sus resultados de salud [9].

Con respecto a los videojuegos relacionados a la mejora de la salud, encontramos aquellos cuya finalidad no es sólo enseñarles a los pacientes los procesos que se están llevando a cabo dentro de su cuerpo, sino también reducir el estrés o la ansiedad derivados de los tratamientos. Estos videojuegos tratan de transmitir la desmitificación de las enfermedades, convirtiéndolas en algo más natural y menos monstruoso de lo que a menudo nos parecen, logrando de esta manera una mejor aceptación de la enfermedad por parte de los pacientes. Algunos que podemos mencionar son:

- Re-Mission [10], fue creado específicamente para jóvenes con cáncer infantil. Los jugadores pilotan a Roxxi, un nanobot que viaja a través de los cuerpos de los pacientes con cáncer para destruir las células cancerosas, luchando contra las infecciones bacterianas y la gestión de efectos secundarios derivados del cáncer y su tratamiento.
- Zak vs Mutantes, el protagonista del juego, Zak y tiene diabetes. Durante la

aventura que vivirá Zak, necesitará vigilar sus niveles de glucosa en sangre, comer para recuperar energías cuando lo necesite, hacer un poco de ejercicio y encontrar a los mutantes que invaden la ciudad. El juego pretende que tanto los chicos con diabetes como sus familias sean conscientes de que pueden vivir y jugar como los demás chicos si se hacen responsables de su enfermedad [11].

Dentro de los videojuegos relacionados a la prevención y promoción de la salud, encontramos aquellos que nos acercan a conductas y conocimientos sobre salud:

- DigesTower, tiene como escenario el sistema digestivo humano. El juego está enfocado para niños en edad escolar. El personaje se llama Elise. El objetivo principal del juego es comprender la importancia de una alimentación saludable [6].
- Virtual Drug Use, intenta hacer conscientes a los jóvenes de los problemas de la adicción a las drogas, en concreto a la cocaína. El juego permite la opción de correr una carrera de motos, con drogas o sin drogas. Los investigadores propusieron el siguiente modo de juego: primero jugar una carrera tomando drogas. La ventaja de este modo es que eres casi invencible (no te puedes caer de la moto, saltas a enorme altura, no hay sensación de peligro) pero cada poco tiempo hay que parar a tomar una dosis, lo que interrumpe el juego dificultando la carrera. La segunda partida será sin consumir drogas, lo que permite al jugador medir su propia habilidad, ya que el motociclista se puede caer de la moto y no conduce tan deprisa, pero el juego es continuo [12].
- Pamoja Mtaani, centra sus esfuerzos en la prevención y control de enfermedades de transmisión sexual. Pamoja Mtaani (Juntos en el barrio) es un videojuego multijugador para PC basado en un

mundo 3D abierto para jugar on-line. Contiene mensajes de prevención sobre el virus del SIDA que pretenden cambiar el comportamiento de los jóvenes frente a esta enfermedad [13].

Una característica de los juegos mencionados anteriormente, es que no están orientados al movimiento del participante sino más bien a un juego pasivo, de movimiento reducido que obliga al jugador a estar sentado para ejecutarlo imposibilitándolo de explorar su entorno y objetos que lo rodean.

El uso de la realidad aumentada en videojuegos adaptados para distintas discapacidades es una herramienta innovadora y eficaz como apoyo del aprendizaje. Permite a los participantes desplazarse interactuando con su entorno y con los objetos reales que lo rodean.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto de investigación Juego de Realidad Aumentada para Incentivar la Actividad Física en Niños con Discapacidad tiene como objetivo general incentivar el hábito de caminata en los niños con diagnósticos dentro espectro autista, trastornos de marcha, diabetes, sobrepeso, obesidad y aquellos en donde sin tener una discapacidad con sintomatología general se recomienda el ejercicio de caminata controlado.

A partir de esta línea de investigación creamos un juego de realidad aumentada, permitiendo la interacción del participante con el mundo real y los objetos que lo rodean, invitándolo a explorarlos e interactuar con ellos mediante el movimiento y el desplazamiento. Esto permitirá a los niños adquirir nuevas experiencias a nivel motor abandonando el sedentarismo que hasta ahora fueron propuestos por otros videojuegos orientados a la salud.

El juego de movimiento ayuda además a desarrollar el autoestima, la autonomía, la capacidad cognitiva, la conciencia de su propio cuerpo y el reconocimiento de la relación cuerpo/espacio/objeto. Las aplicaciones de realidad aumentada permiten ampliar el universo del juego a través de modificaciones virtuales en el mundo real, que sólo ocurren a través del dispositivo con el que se está observando el escenario. De esta forma, además de jugar en el mundo real, se incentiva el movimiento por lo que los juegos dejan de ser sedentarios; la interacción con el juego se realiza en el mismo mundo donde el participante se mueve habitualmente.

Un aspecto importante de esta línea de investigación es la implementación de herramientas de acceso para las diferentes necesidades de los jugadores. El juego tiene un menú adaptado donde cada usuario podrá optar de acuerdo a sus necesidades y/o impedimentos, el menú que más se adecue a sus capacidades, destacando un lector de voz y la posibilidad de modificar los textos a mayúscula para participantes disminuidos visuales o que no poseen lectoescritura. También se destaca el control parental, que es una característica especialmente útil para padres, delimitando el tiempo de uso del juego por parte del participante para que el mismo sea consumido en forma óptima y regulada.

Cada juego define su propio escenario, donde se puede restringir el área del mismo a determinada cantidad de metros alrededor de la posición actual, sonando una alarma al salirse del perímetro establecido. Por otra parte, permite tener abierta múltiples sesiones o juegos, de tal forma que si el usuario deja de jugar o suspende el juego, cuando desee y esté en la zona del escenario puede retomarlo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Con esta línea de investigación pretendemos:

- Incentivar el hábito de la caminata en los niños con discapacidad.
- Promover el desarrollo de la actividad física a través del uso de juegos de realidad aumentada.
- Fortalecer la importancia del movimiento para el desarrollo físico y mental de los niños.
- Generar un ambiente de aprendizaje que les permita a los niños adquirir nuevos conocimientos a través del uso del juego.

Actualmente desarrollamos “AXxion: Cofre del Tesoro Perdido”. Un app de realidad aumentada compatible con el sistema operativo Android. Este juego se basa en la búsqueda del tesoro, en donde el jugador seguirá pistas que le permitan encontrar los tesoros olvidados por un pirata y también deberá responder preguntas para poder avanzar de nivel.

La implementación del proyecto se realizó en el lenguaje de programación Java, utilizando como entorno de desarrollo Eclipse, motor de bases de datos SQLite y la herramienta Android SDK de Google.

El juego se basa en un motor de diseño propio y permite administrar servicios del equipo tales como posicionamiento preciso, cámara, acelerómetro, giróscopo, y compás; siempre orientado a juegos de realidad aumentada. Se aplica un conjunto de filtros y funciones de tratamiento de señales para trabajar con señales estables y que representen la realidad. Brinda un conjunto de servicios que facilitan la construcción del juego, permitiendo mantener objetos geo-localizados, el cálculo de ángulos, posiciones y tamaños en los que se verían dichos objetos de acuerdo a los 6 grados de libertad con los que cuenta el dispositivo móvil. Este motor de juego ha sido desarrollado totalmente a medida para este tipo de aplicaciones.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de esta línea de trabajo consiste en cinco personas. El Dr. Nelson Acosta como director. El Ing. Alejandro Cruz, Ingeniero en Sistemas de la Universidad Nacional de Chilecito (UNDeC), actualmente realizando un Doctorado en la UNICEN con beca doctoral de la Agencia. También contamos con la participación a tiempo parcial de tres alumnos de Ingeniería en Sistemas.

5. BIBLIOGRAFÍA

- [1] Speiser PW, Rudolf MC, Anhalt H, Camacho-Hubner C, Chiarelli F, Eliakim A, Freemark M, Gruters A, Herskovitz E, Iughetti L *et al.* Childhood obesity. *J Clin Endocrinol Metab* 2005, 90:1871-1887.
- [2] Livingstone MB. Childhood obesity in Europe: a growing concern. *Public Health Nutr* 2001, 4:109-116.
- [3] Caracuel J. C, Arbinaga F. Repercusión del ejercicio físico sobre la salud. Apuntes de Psicología. 2012, Vol. 30 (1-3), págs. 547-554
- [4] Casajús J. A, Vicente-Rodriguez G. Ejercicio físico y salud en poblaciones especiales. Exernet, 2011.
- [5] Roncancio Melgarejo C. R, Sichacá Ávila E. G, La actividad Física Como Juego en la Educación Inicial de los Niños Preescolares, Trabajo de Grado: Especialistas en Epidemiología, Universidad de Antioquia - Facultad Nacional De Salud Pública Hector Abad Gómez, 2009.
- [6] Dias JD, Mekaro MS, Lu JKC, Otsuka JL, Fonseca LMM, Zem-Mascarenhas SH. Serious game development as a strategy for health promotion and tackling childhood obesity. *Rev.Latino-Am. Enfermagem.* 2016;24:e2759. DOI: <http://dx.doi.org/10.1590/1518-8345.1015.2759>
- [7] Tejeiro, S. y Pelegrina del Río, M. (2008). La Psicología de los videojuegos. Un modelo de investigación. Málaga. Ediciones Aljibe.
- [8] Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining “gamification”. Proceedings of the 15th International Academic MindTrek Conference, September 28-30, Tampere, Finland; 2011.
- [9] Kato, P. M. (2010). Video games in health care: Closing the gap. *Review of General Psychology* 2010, Vol. 14, No. 2, 113–121.
- [10] Kato, P. M., Cole, S. W., Bradlyn, A. S., & Pollock, B. H. (2008). A video game improves behavioral outcomes in adolescents and young adults with cancer: A randomized trial. *Pediatrics*, 122, Number 2.
- [11] Buena Salud. "Un juego de video para niños con Diabetes", Octubre 2010. <http://www.revistabuenasalud.cl/un-juego-de-video-para-ninos-con-diabetes/>
- [12] Adicción Drogas, Prevención de la Adicción a las Drogas. <https://www.tavad.com/juego-adiccion-virtual/>
- [13] Virtual Heroes, 2008. <http://www.virtualheroes.com/portfolio/Our%20Portfolio/Pamoja-Mtaani>

Algoritmos de Layout Automático para una Herramienta Multi-Vistas de Modelado Ontológico

Giuliano Marinelli¹ Germán Braun^{1,2,3} Laura Cecchi¹
Pablo Fillottrani^{2,4}

email: giuliano.marinelli@est.fi.uncoma.edu.ar,
{german.braun,lcecchi}@fi.uncoma.edu.ar, prf@cs.uns.edu.ar

¹*Grupo de Investigación en Lenguajes e Inteligencia Artificial*
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

²*Laboratorio de I&D en Ingeniería de Software y Sistemas de Información*
Departamento de Ciencias e Ingeniería de la Computación
UNIVERSIDAD NACIONAL DEL SUR

³*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*

⁴*Comisión de Investigaciones Científicas de la provincia de Buenos Aires (CIC)*

Resumen

La presente investigación se desarrolla mediante el trabajo de docentes investigadores de la Universidad Nacional del Comahue (UNCo) y de la Universidad Nacional del Sur (UNS), en el contexto de proyectos de investigación financiados por las universidades indicadas.

La complejidad inherente a los modelos conceptuales y a las ontologías requieren mantener la legibilidad de dichos modelos. Para cubrir esta necesidad y, teniendo en cuenta el tamaño de los diagramas en términos de sus primitivas gráficas, algoritmos efectivos y eficientes de *layout* deben ser desarrollados para garantizar la usabilidad de las herramientas.

El objetivo de esta línea de investigación es analizar, diseñar e implementar algoritmos de *layout* en una herramienta gráfica Web. Para esto, se profundizará en técnicas y heurísticas de visualización basadas en la teoría de *Crossing Number* y sus variantes, aplicado a los lenguajes EER, UML y ORM 2.

Palabras Clave: Ingeniería de Software basada en Conocimiento, Ontologías, Interoperabilidad

de Lenguajes de Modelado Conceptual.

Contexto

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el marco del proyecto de investigación *Agentes Inteligentes y Web Semántica (04/F014)*, por la Universidad Nacional del Sur a través del proyecto de investigación *Integración de Información y Servicios en la Web (24/N027)* y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral. Los proyectos de investigación tienen una duración de cuatro años y la beca doctoral una duración de 5 años, finalizando esta última en abril de 2019.

1. Introducción

En la actualidad disponemos de herramientas que asisten a los modeladores para la realización de tareas de modelado conceptual y ontológico, las cuales brindan un soporte gráfico de manera que se simplifique la visualiza-

ción del contenido y facilite dicha tarea. Esto brinda al modelador mayor rapidez en los procesos cognitivos que realiza a la hora de modelar, dándole la posibilidad de tomar mejores decisiones de diseño. El modelado implica el desarrollo de diagramas a través de lenguajes estándar como UML [4], ER [8], ORM 2 [16], entre otros, los cuales son, por lo general, soportados por las herramientas relevadas en la literatura [1, 2].

Trabajar con tales herramientas siempre presenta alguna dificultad (o poca facilidad), particularmente, cuando se modelan dominios de aplicación reales, donde la cantidad de elementos gráficos es considerable o cuando se realiza la integración de múltiples modelos. Estos problemas son los relacionados con la visualización o *layout* [9, 10], como la distribución de los elementos gráficos en el espacio del diagrama, o las formas y tamaños de estos elementos.

Herramientas para modelado ontológico tales como Protégé [18], OWLGrEd [7], VOWL [19] e ICOM [11], poseen algoritmos de *layout* automático, pero con ciertas limitaciones en relación al enfoque de nuestra propuesta. En primer lugar, OWLViz, OntoGraf y SOVA son plugins de Protégé que utilizan grafos como lenguaje gráfico para ontologías. Por lo tanto, si bien ofrecen algoritmos visuales, están limitados en cuanto a la expresividad gráfica subyacente. ICOM y OWLGrEd están basados en los lenguajes de modelado conceptual ER y UML, respectivamente. Ambos permiten representar ontologías OWL [22] y poseen capacidades de *layout* automático. Sin embargo, solo proveen un conjunto de alineamientos fijos y estáticos. Además, el soporte gráfico de ICOM está actualmente discontinuado. Por último, VOWL también implementa una visualización dinámica pero, al igual que los plug-ins previos, usa grafos como modelos para OWL y no lenguajes de modelado conceptual.

Existen otras dificultades además de la visualización, a la hora de plantear soluciones a problemas de dominio, donde se requiere un análisis semántico de los elementos y las relaciones que lo definen. Esta es una de las principales motivaciones de nuestra herramienta, actualmente en desarrollo, *crowd* [14, 15, 5]. *crowd*

da soporte gráfico al modelador en el diseño de modelos conceptuales y ontologías con razonamiento automático [3] para validarlos, lo que facilita la creación y depuración de modelos basados en un dominio específico. La herramienta también provee soporte gráfico para los lenguajes ER, UML y ORM 2, junto con un metamodelo [12], posibilitando la interoperabilidad entre ellos.

En esta línea de investigación proponemos trabajar sobre una mejora en los algoritmos de visualización de los diagramas, y sobre la resolución de los problemas que esto plantea. En este sentido, utilizaremos la herramienta *crowd* para la implementación de nuevos prototipos y simultáneamente este trabajo permitirá la ampliación de la herramienta.

Las problemáticas principales a abarcar tienen un motivo estético sobre la visualización de diagramas (más específicamente sobre grafos) y a la vez subyacen en problemas de mayor índole. Entre los problemas encontramos el número de cruzamientos entre los arcos que conectan los elementos de un grafo, conocido como *Crossing Number* [13], que consiste en obtener el mínimo número de cruces entre arcos posible mediante la reorganización de los elementos del grafo y el trazado de sus arcos. Tal problema se clasifica como NP-Completo y, por ello, no se concibe una solución eficiente para este. Éste y otros aspectos estéticos, como la distancia óptima entre los elementos de un grafo, que el grafo ocupe demasiado espacio, y la estructura de las líneas que representan los arcos, entre otros, son objetivos de estudio.

crowd no dispone de algoritmos de *layout*, por lo que es de importancia el desarrollo y estudio de los mismos, de manera que se pueda complementar la herramienta y cumplir en mayor medida su objetivo, que en sí mismo es asistir a los modeladores en el proceso de desarrollo de modelos conceptuales y ontologías. Para ello se implementarán soluciones prácticas sobre la herramienta para mejorar la visualización de los diagramas en la misma, y por otro lado esto permitirá estudiar y adentrarse en mayor medida en la resolución de problemas NP.

La estructura de este trabajo es la siguiente.

A continuación se presenta la línea de investigación y los proyectos en los que se enmarca. En la sección 3, se describen los resultados que se esperan de este trabajo. Finalmente, se detalla la formación en recursos humanos en la temática de la línea de investigación.

2. Líneas de investigación y desarrollo

El proyecto de investigación Agentes Inteligentes y Web Semántica, UNCo, tiene varios objetivos generales. Uno de ellos es el de generar conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la web, es decir lo que se ha llamado la Web Semántica. En este sentido, se estudian técnicas de representación de conocimiento y razonamiento, metodologías de modelado conceptual y mecanismos para la interoperabilidad de aplicaciones, tanto a nivel de procesos como de datos. Fundamentalmente, se busca aplicar estos conceptos como soporte para comunidades de desarrollo de ontologías.

Por otro lado, en el proyecto de investigación Integración de Información y Servicios en la Web se propone investigar y desarrollar metodologías y herramientas que favorezcan la interoperabilidad semántica de información y de servicios en la Web, fundamentados en los últimos avances en el área de lenguajes de representación del conocimiento, ontologías y modelado conceptual.

Ambos proyectos confluyen en la línea de investigación de este trabajo, en la que se explora entre otros, sobre temas afines a la Ingeniería de Software basada en Conocimiento y la Ingeniería de Conocimiento. Particularmente, se ha escogido experimentar sobre metodologías que integren razonamiento con un front-end gráfico para dar soporte a la ingeniería de ontologías.

Concretamente, esta línea de investigación tiene como objetivo principal el desarrollo de algoritmos de *layout* automático junto con la implementación de un prototipo del mismo sobre la arquitectura de *crowd*, con el propósito de

expandir las funcionalidades de la herramienta. Con tal idea, la planificación consta de varias etapas. En un principio se realizará un relevamiento de las tecnologías y/o investigaciones relevantes [21, 17], de manera que se pueda adquirir los conocimientos necesarios para lograr el objetivo, tanto de manera teórica, como es el análisis de problemas subyacentes como *Crossing Number* y *Optimal Linear Arrangement*, entre otros posibles, y por otro lado, el estudio de algoritmos y soluciones buscando integrar una nueva solución adaptada a las necesidades de *crowd*.

crowd busca implementar un metamodelo [12] que permita trabajar con múltiples lenguajes (UML, ER y ORM 2 en un principio) y que puedan interoperar. En este sentido, el algoritmo de *layout* a desarrollar se centrará en un diagrama general, es decir, un grafo con elementos uniformes, de manera de aplicar técnicas y optimizaciones de visualización en tal grafo sin considerar las especificidad de cada diagrama en concreto. A continuación, se tratará cada diagrama de manera específica teniendo en cuenta las características propias de los mismos, sus primitivas gráficas y un relevamiento de modelos conceptuales y ontológicos reales. Asimismo, se espera integrar aspectos de visualización con aspectos semánticos propios de los modelos [20].

3. Resultados esperados

El objetivo general de este trabajo es el desarrollo de un algoritmo genérico de *layout* automático para la herramienta de soporte a la ingeniería ontológica: *crowd* y así empoderar el proceso de visualización definido para esta herramienta por los autores de este trabajo [6]. Se pretende trabajar en un diseño que permita adaptar las funcionalidades de visualización de las ontologías, considerando los diferentes lenguajes de modelado conceptual, como ER, UML y ORM 2.

Se espera que los expertos o modeladores puedan utilizar la herramienta de acuerdo a sus preferencias de notación con una disposición

apropiada de los elementos visuales de los lenguajes de modelado conceptual, lo que redituará en una mejor legibilidad y un adecuado entendimiento del diagrama, potenciando la comunicación entre los interesados en el conocimiento de los modelos conceptuales y las ontologías.

4. Formación de recursos humanos

Sobre la temática de esta línea de investigación, uno de los autores de este trabajo está desarrollando su tesis de grado de la Licenciatura en Ciencias de la Computación.

Por otra parte, otro de los autores de este trabajo está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur y (beca interna doctoral CONICET).

Bibliografía

- [1] Martin Auer, Ludwig Meyer, and Stefan Biffl. Explorative uml modeling-comparing the usability of uml tools. In *ICEIS (3)*, pages 466–474, 2007.
- [2] Martin Auer, Thomas Tschurtschenthaler, and Stefan Biffl. *A flyweight uml modeling tool for software development in heterogeneous environments*. IEEE, 2003.
- [3] Franz Baader. *The description logic handbook: Theory, implementation and applications*. Cambridge university press, 2003.
- [4] Grady Booch. *The unified modeling language user guide*. Pearson Education India, 2005.
- [5] Germán Braun, Christian Gimenez, Laura Cecchi, and Pablo Fillottrani. Towards a visualisation process for ontology-based conceptual modelling. In *ONTOBRAS–Brazilian Ontology Research Seminar (ONTOBRAS)*, number 1862.
- [6] Germán Braun, Christian Gimenez, Laura Cecchi, and Pablo Fillottrani. Towards a Visualisation Process for Ontology-Based Conceptual Modelling. In *ONTOBRAS*, 2016.
- [7] K Cerans, J Ovcinnikova, Renars Liepins, and A Sprogis. Advanced owl 2.0 ontology visualization in owlged. In *Databases and Information Systems VII–Selected Papers from the 10th Int. Baltic Conference (DB&IS)*, 2012.
- [8] Peter Pin-Shan Chen. The entity-relationship model—toward a unified view of data. In *Readings in artificial intelligence and databases*, pages 9–36. Elsevier, 1988.
- [9] Holger Eichelberger. On class diagrams, crossings and metrics. In *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2006.
- [10] Holger Eichelberger and Jurgen Wolff von Gudenberg. Uml class diagrams-state of the art in layout techniques. In *Proceeding of Vissoft 2003, International Workshop on Visualizing Software for Understanding and Analysis*. Citeseer, 2003.
- [11] Pablo R Fillottrani, Enrico Franconi, and Sergio Tessaris. The icom 3.0 intelligent conceptual modelling tool and methodology. *Semantic Web*, 3(3), 2012.
- [12] Pablo Rubén Fillottrani and C Maria Keet. Conceptual model interoperability: a metamodel-driven approach. In *International Workshop on Rules and Rule Markup Languages for the Semantic Web*, pages 52–66. Springer, 2014.
- [13] Michael R Garey and David S Johnson. Crossing number is np-complete. *SIAM Journal on Algebraic Discrete Methods*, 4(3), 1983.
- [14] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Rubén Fillottrani. crowd: A tool for conceptual modelling assisted by automated reasoning. In *II*

- Simposio Argentino de Ontologías y sus Aplicaciones (SAOA 2016)-JAIIO 45 (Tres de Febrero, 2016)*, 2016.
- [15] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Rubén Fillotrani. Interoperabilidad entre lenguajes de modelado conceptual en crowd. In *XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires)*, 2017.
- [16] Terry Halpin. Orm 2. In *OTM Confederated International Conferences On the Move to Meaningful Internet Systems*. Springer, 2005.
- [17] Hao Hu, Jun Fang, Zhengcai Lu, Fengfei Zhao, and Zheng Qin. Rank-directed layout of uml class diagrams. In *Proceedings of the First International Workshop on Software Mining*, pages 25–31. ACM, 2012.
- [18] Holger Knublauch, Ray W Ferguson, Natalya F Noy, and Mark A Musen. The protégé owl plugin: An open development environment for semantic web applications. In *International Semantic Web Conference*. Springer, 2004.
- [19] Steffen Lohmann, Stefan Negru, and David Bold. The protégéowl plugin: ontology visualization for everyone. In *European Semantic Web Conference*. Springer, 2014.
- [20] Reyhaneh Raissi. Improving readability of software model with adding visualized extra information to uml diagram. In *Proceedings of the International Conference on Compute and Data Analysis*, pages 28–32. ACM, 2017.
- [21] Dabo Sun and Kenny Wong. On evaluating the layout of uml class diagrams for program comprehension. In *Program Comprehension, 2005. IWPC 2005. Proceedings. 13th International Workshop on*, pages 317–326. IEEE, 2005.
- [22] World Wide Web Consortium (W3C). OWL 2 Web Ontology Language Document Overview (Second Edition), 2012. <http://www.w3.org/TR/owl2-overview/>, accedido en Junio de 2013.

Hacia un Lenguaje Gráfico para SPARQL-DL

Christian Gimenez¹

Germán Braun^{1,2,3}

Laura Cecchi¹

Pablo Fillottrani^{2,4}

email: {christian.gimenez, german.braun, lcecchi}@fi.uncoma.edu.ar,
prf@cs.uns.edu.ar

¹*Grupo de Investigación en Lenguajes e Inteligencia Artificial*
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

²*Laboratorio de I&D en Ingeniería de Software y Sistemas de Información*
Departamento de Ciencias e Ingeniería de la Computación
UNIVERSIDAD NACIONAL DEL SUR

³*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*

⁴*Comisión de Investigaciones Científicas de la provincia de Buenos Aires (CIC)*

Resumen

Esta línea de investigación se desarrolla en forma colaborativa entre docentes-investigadores de la Universidad Nacional del Comahue y de la Universidad Nacional del Sur, en el marco de proyectos de investigación financiados por las universidades antes mencionadas.

El objetivo general de este trabajo es el desarrollo de metodologías que integren consultas textuales y gráficas con razonamiento automático, en ambientes de modelado conceptual y ontológico. Así como el uso de razonamiento para la identificación de errores en la ingeniería ontológica es vital y tiene consecuencias evidentes, es esencial que sea también considerado en la generación de consultas, junto con el soporte gráfico para la interacción de los usuarios con los modelos.

Palabras Clave: Consultas gráficas, Ingeniería de Software basada en Conocimiento, Lógicas Descriptivas, Ontologías.

Contexto

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el

marco del proyecto de investigación Agentes Inteligentes y Web Semántica (04/F014) y de una Beca de Investigación de en la categoría Graduado de Iniciación; por la Universidad Nacional del Sur a través del proyecto de investigación Integración de Información y Servicios en la Web (24/N027) y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral. Los proyectos de investigación tienen una duración de cuatro años y la beca doctoral una duración de 5 años, finalizando esta última en abril de 2019. La Beca de Investigación UNCo en la categoría Graduado de Iniciación tiene una duración de un año.

1. Introducción

En trabajos previos [1, 2, 3, 4] se ha presentado a *crowd*, una herramienta para el modelado conceptual y ontológico. El objetivo es el de permitir a los usuarios realizar modelos de forma visual y colaborativa, con asistencia de servicios de razonamiento para determinar inconsistencias. Esto es posible debido a la formalización en Lógica Descriptiva propuesta por [5],

la cual puede ser expresada de forma textual y consultada por consistencia utilizando razonadores como Racer [6] o Koncude [7].

Se ha mostrado que aplicar técnicas visuales para el acceso de datos es particularmente exitoso. Esto se debe al incremento de usuarios no expertos que no están familiarizados con lenguajes de consultas de bases de datos como SQL. Además, en comparación con el uso del lenguaje natural, las consultas visuales superan el problema de depender en el lenguaje del usuario y las limitaciones impuestas por el área de aplicación [8]. Es importante proveer a los expertos de dominios con herramientas para la formulación de consultas para expresar sus necesidades de información en términos de consultas sobre las ontologías [9].

Las ontologías, ya sean modeladas gráficamente o incorporadas textualmente utilizando alguna sintaxis formal (por ejemplo, OWL 2 [10]), pueden ser consultadas en cuanto a su nivel conceptual para poder estudiarlas o para hacer consultas que integren tanto el nivel terminológico (TBox), el de relaciones (RBox) como el de aserciones (ABox). Esto se puede llevar a cabo por medio de consultas SPARQL-DL [11, 12]. Este lenguaje de consulta se define para OWL-DL y puede combinar consultas del TBox, RBox y ABox, a diferencia de DIG [13] y nRQL [14] que están limitados a consultas atómicas. Además, se alinea con SPARQL [15] para mejorar la interoperabilidad de aplicaciones con la Web Semántica.

Sin embargo, basados en las ventajas de las consultas visuales expresadas en los párrafos anteriores, es posible expresarlas por medio de primitivas visuales.

La estructura del presente trabajo es la siguiente. En la sección 2 presentamos los objetivos de los proyectos de investigación en los que se enmarca este trabajo y describimos la línea de investigación actual. En la sección 3 indicamos algunos resultados obtenidos y trabajos futuros. Finalmente, comentamos aspectos referentes a la formación de recursos humanos en esta temática.

2. Línea de Investigación y Desarrollo

El proyecto de investigación Agentes Inteligentes y Web Semántica, UNCo, tiene varios objetivos generales. Uno de ellos es el de generar conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en la Web Semántica, para entre otros, dar soporte a comunidades de desarrollo de ontologías. En este sentido, y compartido con el proyecto de investigación Integración de Información y Servicios en la Web, de la UNS, se estudian técnicas de representación de conocimiento y razonamiento, metodologías de modelado conceptual y mecanismos para la interoperabilidad de aplicaciones, tanto a nivel de procesos como de datos.

Ambos proyectos confluyen en la línea de investigación de este trabajo, en la que se explora entre otros, sobre temas afines a la representación del conocimiento, las Lógicas Descriptivas [16], las Ontologías, la Ingeniería de Software basada en Conocimiento y la Ingeniería de Conocimiento.

En esta línea de investigación se propone desarrollar un lenguaje gráfico que represente las consultas SPARQL-DL a realizarse sobre una ontología. Inicialmente, se describirán primitivas visuales para consultar las estructuras del TBox. Las consultas visuales se traducirán a su versión textual, las cuales en conjunto con una ontología serán ejecutadas sobre un motor SPARQL-DL para obtener las respuestas. Este proceso se describe en el siguiente párrafo con un ejemplo.

Suponemos un modelo extraído del ejemplo de [5] de un diagrama UML mostrado en la Figura 1. Se realiza una consulta visual con la que se desea buscar las clases "Y" que sean subclases de "X". Ésta se transforma en una consulta SPARQL-DL la cual se utilizará para alimentar al motor junto con la formalización del diagrama UML a DL. Las respuestas obtenidas, son mostradas al usuario en forma textual, pudiendo ser interpretadas gráficamente reemplazando las variables por los nombres de clases encontrados.

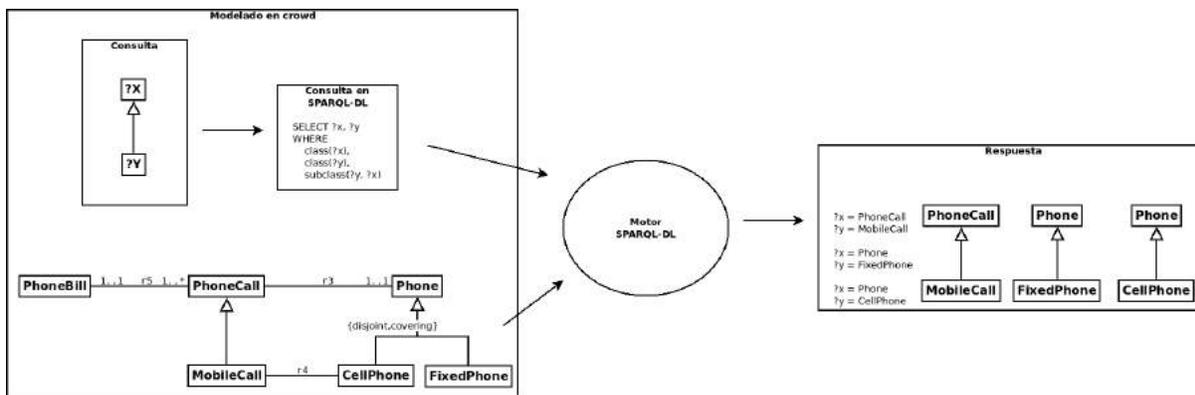


Figura 1: Ejemplo de consulta visual.

Para implementar este lenguaje, se continuará con el desarrollo de la herramienta Web denominada *crowd*, la cual contribuye al modelador en el diseño y la visualización de ontologías durante su evolución, actualmente por medio del lenguaje UML, asistidos por técnicas de razonamiento automáticas. Para ello, se implementará una extensión de la arquitectura que consistirá en un diseño de interfaz que permita el modelado de consultas utilizando dicho lenguaje gráfico. En el servidor, se llevará a cabo la traducción y su ejecución sobre una ontología ofrecida por el usuario, mostrando los resultados en el cliente.

3. Resultados Obtenidos y Trabajo Futuro

Actualmente, *crowd* permite crear modelos basados en un subconjunto de primitivas de UML. Puede determinar las inconsistencias del modelo debido a la formalización a Lógica Descriptiva del modelo basándonos en la propuesta en [5], y posteriormente utilizando un programa razonador. También, es posible obtener la representación textual del modelo formalizado en OWL 2 [17, 10] para ser utilizado con otras herramientas. Un prototipo está disponible en la URL crowd.fi.uncoma.edu.ar.

Además, está en desarrollo el soporte al metamodelo descrito en [18] para brindar interoperabilidad entre los lenguajes UML, EER y ORM considerando sus diferencias semánticas. Esto

implica la necesidad de incorporar las primitivas gráficas correspondientes a los lenguajes indicados y la implementación del metamodelo en el servidor.

Se propone definir un lenguaje visual para la realización de consultas sobre las ontologías modeladas. Dicho lenguaje se pretende que posea primitivas que sean similares a los lenguajes de modelado ampliamente usados en la ingeniería de Software para minimizar la curva de aprendizaje. De esta manera, se espera poder buscar patrones que se presenten en la ontología mezclando consultas de nivel conceptual (TBox) como del nivel de datos (ABox).

Finalmente, como trabajo futuro a este desarrollo se considera implementar una extensión de *crowd* que permita al usuario el modelado de las consultas y la ontología, como así también la importación de los datos necesarios para obtener las respuestas.

4. Formación de Recursos Humanos

Uno de los autores de este trabajo está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur (beca interna doctoral CONICET).

En la Universidad Nacional del Comahue, Facultad de Informática, se otorgaron Becas de Investigación en la categorías de Graduado de Iniciación. Una de esas becas fue otorgada a uno de los autores de este trabajo. En el marco de es-

ta beca, se espera que el investigador comience un posgrado.

Referencias

- [1] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Fillottrani. Una Arquitectura Cliente-Servidor para Modelado Conceptual Asistido por Razonamiento Automático. In *XVIII Workshop de Investigadores en Ciencias de la Computación*, 2016.
- [2] Christian Gimenez, Germán Braun, Laura Cecchi, and Laura Fillottrani. crowd: A Tool for Conceptual Modelling assisted by Automated Reasoning - Preliminary Report. In *the 2nd Simposio Argentino de Ontologías y sus Aplicaciones SAOA '16 JAIIO '16*, 2016.
- [3]
- [4] Germán Braun, Christian Gimenez, Pablo Fillottrani, and Laura Cecchi. Towards conceptual modelling interoperability in a web tool for ontology engineering. In *the 3rd Simposio Argentino de Ontologías y sus Aplicaciones SAOA '17 JAIIO '17*, 2017.
- [5] Daniela Berardi, Diego Calvanese, and Giuseppe De Giacomo. Reasoning on UML class diagrams. *Artif. Intell.*, 168(1-2):70–118, 2005.
- [6] V. Haarslev and R. Möller. Racer system description. In R. Goré, A. Leitsch, and T. Nipkow, editors, *International Joint Conference on Automated Reasoning, IJ-CAR'2001, June 18-23, Siena, Italy*, pages 701–705. Springer-Verlag, 2001.
- [7] Andreas Steigmiller, Thorsten Liebig, and Birte Glimm. Konclude: System description. *Journal of Web Semantics (JWS)*, 27:78–85, 2014.
- [8] Amineh Fadhil. Ontovql: Ontology visual query language. Master's thesis, Concordia University Montreal, Quebec, Canada, September 2008.
- [9] Ahmet Soylu, Evgeny Kharlamov, Dmitriy Zheleznyakov, Ernesto Jiménez-Ruiz, Martin Giese, and Ian Horrocks. OptiqueVQS: Ontology-Based Visual Querying. In *Proceedings of the International Workshop on Visualizations and User Interfaces for Ontologies and Linked Data co-located with 14th ISWC*, 2015.
- [10] World Wide Web Consortium (W3C). OWL 2 Web Ontology Language Document Overview (Second Edition), 2012. <http://www.w3.org/TR/owl2-overview/>, accedido en Junio de 2013.
- [11] Ilianna Kollia, Birte Glimm, and Ian Horrocks. Sparql query answering over owl ontologies. In Grigoris Antoniou, Marko Grobelnik, Elena Simperl, Bijan Parsia, Dimitris Plexousakis, Pieter De Leenheer, and Jeff Pan, editors, *The Semantic Web: Research and Applications*, pages 382–396, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.
- [12] Evren Sirin and Bijan Parsia. Sparql-dl: Sparql query for owl-dl. In *In 3rd OWL Experiences and Directions Workshop (OWLED-2007)*, 2007.
- [13] S. Bechhofer, R. Moller, and P. Crowther. The DIG Description Logic Interface. In *In Proc. of International Workshop on Description Logics (DL2003)*, 2003.
- [14] Volker Haarslev, Ralf Möller, and Michael Wessel. Querying the semantic web with racer + nrql. In *In Proceedings of the KI-2004 International Workshop on Applications of Description Logics (ADL'04)*, 2004.
- [15] W3C OWL Working Group. *SPARQL 1.1 Query Language*. W3C Recommendation, 21 March 2013. Available at <http://www.w3.org/TR/sparql11-query/>.

- [16] Diego Calvanese, Maurizio Lenzerini, and Daniele Nardi. Description logics for conceptual data modeling. In *Logics for Databases and Information Systems*, pages 229–263. Kluwer, 1998.
- [17] Thorsten Liebig, Marko Luther, Mariano Rodriguez, Diego Calvanese, Michael Wessel, Ralf Möller, Matthew Horridge, Sean Bechhofer, Dmitry Tsarkov, and Evren Sirin. OWLlink: DIG for OWL 2. In *In Proceedings of the 5th OWL Experiences and Directions Workshop (OWLED-2008)*, page 48, 2008.
- [18] Pablo R. Fillottrani and C. Maria Keet. KF metamodel formalization. *CoRR*, abs/1412.6545, 2014.

Auditoría automatizada de Trazabilidad Vitivinícola

Alejandro Vazquez^{1,3}, Carlos Troglia¹, Carlos Martinez¹, Gustavo Manino¹, Susana Hubbe², Alberto Cortez^{1,3}, Javier Caballero¹, Agustín Espinosa¹, Ariel Martín¹, Mariano García¹, Claudia Naveda^{1,3}.

UTN Facultad Regional Mendoza, Ingeniería en Sistemas de Información, ¹Laboratorio de Auditoría y Seguridad de TIC, ²Maestría en Ingeniería en Calidad, ³Universidad del Aconcagua, Licenciatura en Informática y Desarrollo de Software

avazquez@frm.utn.edu.ar, troglia2000@hotmail.com, carlos.martinez@frm.utn.edu.ar, gfmantino, su_hubbe}@yahoo.com.ar, {cortezalberto, caballerojavier13, agustin.espinosa.21, arielmartin9, marianogarciaamore, claudialaboral}@gmail.com

RESUMEN

La optimización de los mecanismos de auditoría de trazabilidad de vinos es una preocupación creciente en la industria. Entre otros aspectos facilita la aplicación procesos de mejora continua y reingeniería de procesos. Como solución para lograr dicha optimización se presenta en este trabajo los componentes de la arquitectura diseñada para realizar la auditoría de los sistemas de trazabilidad de vinos. Estos componentes deben permitir automatizar las tareas de auditoría y reingeniería, de manera de poder analizar la trazabilidad de los sistemas de manera rápida y plantear acciones para mejorar la calidad en los procesos productivos. En esta investigación se trabaja para obtener una plataforma de auditoría de trazabilidad de vinos. Y como resultado poder aplicar mejora continua y reingeniería de procesos.

CONTEXTO

Este proyecto de investigación se encuentra en ejecución en el marco del laboratorio LabAuSegTIC (Laboratorio de Auditoría y Seguridad de TICs) del Departamento de Ingeniería en Sistemas de Información de la UTN FRM (Universidad Tecnológica Nacional – Facultad Regional Mendoza). Es un proyecto interinstitucional entre UTN FRM y Universidad del Aconcagua, 2016-

2018, homologado por Universidad del Aconcagua mediante Resolución 235/2015 FCSA y Resolución 022/16 CIUDA-Consejo Superior y homologado por UTN mediante el Código de Proyecto Nacional Incentivos EIINIME0003878TC PID 3878.

INTRODUCCIÓN

Esta investigación analiza el contexto y los factores del proceso de producción vitivinícola, para buscar soluciones que mejoren la calidad del producto. Se considera que a partir del estudio, análisis, diseño, y desarrollo de un modelo tecnológico se puede mejorar tanto calidad como trazabilidad del proceso mediante soluciones de hardware-software. La trazabilidad es la capacidad de seguir el recorrido de un alimento a través de las etapas especificadas de producción, procesamiento y distribución [1]. Es considerada como un instrumento de gestión del riesgo. La información brindada mediante trazabilidad permite acotar los alcances de un incidente alimentario, según lo indicado por el Instituto Nacional de Tecnología Industrial [2]. La trazabilidad debe garantizar, además de la trazabilidad interna, que se pueda navegar sobre todos los datos del producto hacia atrás y hacia adelante [3,4]. El sistema debe asegurar:

Trazabilidad hacia atrás: permite conocer las materias primas (ingredientes) que forman parte de un producto, envases y otros

materiales utilizados, así como identificar a sus proveedores.

Trazabilidad hacia delante: permite conocer dónde se ha vendido/distribuido un lote determinado de un producto alimenticio (identificación del producto, lotes, cantidades, fecha de entrega y destinatario).

La trazabilidad interna o del proceso: permite hacer un seguimiento de los productos procesados en el establecimiento y conocer sus características; tratamientos recibidos y circunstancias a las que han estado expuestos.

La plataforma de auditoría de trazabilidad de vinos, debe permitir registrar y rastrear todos estos datos, de forma tal que facilite la toma de decisiones en situaciones de riesgo. Para implementar el rastreo se crearán pistas de auditoría a partir del diseño funcional del sistema, paquetes especializados o el agregado de las pistas posteriormente modificando el sistema.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La línea de investigación propuesta se enfoca en el mejoramiento de las herramientas de auditoría y trazabilidad de vinos que permitan optimizar la gestión del riesgo. El objeto de estudio en este caso es crear un nuevo instrumento para auditoría de trazabilidad de vinos.

Los métodos utilizados en IS tiene similitud con la metodología que se usa en el desarrollo de software. Y se divide en distintas etapas. El tipo de estudio e investigación es "Investigación aplicada con enfoque mixto (cuantitativo y cualitativo) [5,6]. Este enfoque habilita la recolección, análisis y vinculación de datos cuantitativos y cualitativos en un mismo estudio.

La investigación previa y el desarrollo están comprendidos en las siguientes etapas:

1. En la primera etapa se analizaron las leyes y normas de Argentina e internacionales vigentes y pertinentes al proceso productivo del vino y a su trazabilidad.
2. La segunda etapa abarca el estudio, observación y análisis de Sistemas de trazabilidad de vinos existentes.

3. La tercera etapa implica el estudio y análisis de los requerimientos de información de la cadena de producción y comercialización del vino (productor, elaborador, exportador, consumidor). El estudio incluye la identificación y caracterización de fuentes de información relevante en el dominio de aplicación.

4. La cuarta etapa implica la integración de todos los aspectos estudiados para desarrollar una plataforma de auditoría de trazabilidad. El trabajo abarca el análisis, diseño, desarrollo del prototipo y la verificación y validación de sus resultados.

RESULTADOS Y OBJETIVOS

Del análisis previo realizado, según el relevamiento a bodegas de la región de cuyo, se observó que tienen diferentes sistemas de trazabilidad. Algunos con funciones bien integradas, otros con procesos semiautomatizados y otros que sólo registran datos en planillas electrónicas o manualmente, que ayudan parcialmente a determinar la trazabilidad. Por ello, la necesidad de un eficiente modelado de datos sirve como base fundamental para auditar los sistemas y procedimientos [8]. Las posteriores etapas de estudio revelaron la necesidad de crear una Plataforma Open Source de Auditoría de Trazabilidad Genérica utilizando una Arquitectura en Capas. [9 ,10 ,11]. Se estudiaron diversas tecnologías, herramientas y soluciones [12 ,13 ,14 ,15 ,16]. .

Como resultado se definieron algunas ideas guía y soluciones. Así el diseño de los procesos se definió el requerimiento del estándar BPMN 2 (por sus siglas en inglés Business Process Model and Notation) y la necesidad de un motor de procesos de negocio. Además para que las empresas auditadas puedan interoperar con el motor de trazabilidad e implementar procesos de mejora continua se especificó la necesidad de incorporar un API REST (por sus siglas en inglés Representational State Transfer) en la plataforma [22]. Los beneficios de esta propuesta impactan en la mejora de los

procesos productivos y de la calidad del producto.

La arquitectura propuesta está compuesta por:

- Sistema de Gestión de Procesos de Negocio (BPMS por sus siglas en inglés, Business Proceses Management Suites) [17,18].
- Sistema de Gestión de Reglas del Negocio (BRMS por sus siglas en inglés, Business Rules Management Systems) [19].
- Componentes del negocio Java EE [20]
- Plataforma de inteligencia de negocios [21]

En la Figura 1, se muestra el esquema con la arquitectura propuesta.

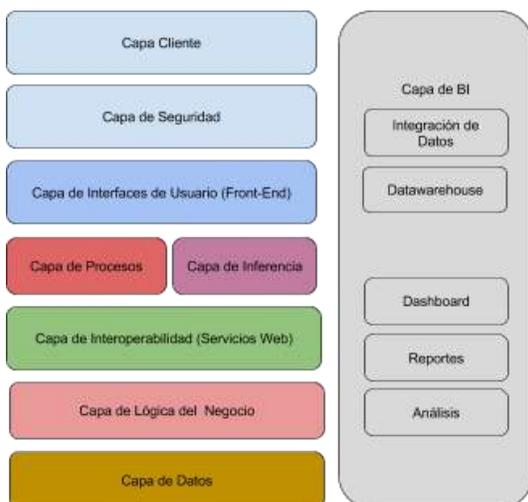


Figura 1 Arquitectura Propuesta

Capa de Cliente, Capa de Seguridad y Capa de Interfaces de Usuario (Front-End)

Los clientes pueden interactuar vía REST [22], con la plataforma de trazabilidad. Para acceder a estos servicios web, previamente se deben autenticar, solicitando autorización y el acceso a los recursos por medio del componente de seguridad. Para la autenticación y/o autorización de usuarios, el estándar recomendado es OpenIDConnect

[23] y Oauth 2 [24]. En este trabajo, se implementa el componente de seguridad con el gestor de acceso e identidad: Keycloak [25]. La interfaz de usuario vinculada a los procesos definidos con BPMN 2 se genera con la herramienta a partir del diseño de una plantilla.

Capa de Procesos

El estándar de la OMG para el modelado de procesos del negocio es BPMN 2.0 (Business Process Model and Notation (BPMN)[21,28]. Para la implementación técnica de BPM se utiliza un tipo de sistemas llamado BPMS (por sus siglas en inglés, Business Process Management System). El BPMS elegido en esta propuesta es BonitaSoft [29], por ser una plataforma Open Source que cumple con requerimientos como: diseñador de formularios de usuarios y fácil interacción con servicios web REST.

Capa de Inferencia

Un sistema de gestión de reglas de negocio (BRMS, por las siglas en inglés de Business Rule Management System) [19]; permite gestionar las reglas del negocio en forma sencilla y eficiente por medio de una interface gráfica. Para la implementación de esta capa se utiliza un sistema de gestión de reglas de negocio, Drools [19]. Es una plataforma: Open Source, completa, refinada y robusta, para la gestión de reglas del negocio.

Capa de Interoperabilidad

La arquitectura orientada a servicios es implementada utilizando REST. Una tarea de servicio de BPMN 2.0, toma los datos cargados en la pantalla relacionada con la tarea anterior y se conecta vía REST al servicio correspondiente para persistir los datos registrados.

Capa de Lógica de Negocios y Capa de Datos

Enterprise Services Bean implementa la lógica del negocio compuesta por: operaciones CRUD (por sus siglas en inglés, Create, Read, Update and Delete) sobre los recursos. Incorporando validaciones y la

interacción con las reglas del negocio. En esta capa se puede trabajar con tecnologías como Java EE, con el despliegue sobre un servidor de aplicaciones como Wilfly [30]. La persistencia se implementa con el framework JPA [31] Hibernate [32], que permite el mapeo objeto/relacional en aplicaciones. La tecnología empleada permite la utilización de distintas tecnologías de Base de datos.

Capa de Inteligencia de Negocios

Para el desarrollo de esta capa se utiliza la metodología Hefesto para la construcción de un Data Warehouse, como se describe en [33]. Esta capa toma los datos almacenados en la estructura de datos y mediante un proceso ETL (por sus siglas en inglés, Extract, Transform and Load), los vuelca en un esquema en estrella.

En base a los datos almacenados en el esquema en estrella se diseña un cubo que se publica en una plataforma de Inteligencia de Negocios como Pentaho [34]. Una vez que el cubo está publicado se diseñan reportes, dashboards y se pueden hacer análisis multidimensionales dinámicos. Por ejemplo, una consulta multidimensional que se puede realizar con este cubo puede ser: "Ponderación por componente de auditoría de trazabilidad, por empresa, por producto, por unidad de tiempo".

proyecto implica estudiar modelos existentes, para determinar las ventajas y desventajas de ellos. Y de esta manera poder plantear un modelo nuevo que supere las limitaciones de las anteriores.

Dentro de los objetivos se definieron generales y específicos:

Objetivos generales

- Diseñar una arquitectura y desarrollar un plataforma que permita auditar los sistemas de trazabilidad y aplicar la reingeniería en la cadena de producción, a los efectos contribuir a la calidad e inocuidad de alimentos.
- Servir de base como herramientas de futuras certificaciones de calidad para los sistemas de trazabilidad de alimentos y construir los componentes informáticos necesarios.

- Desarrollar una plataforma que se ajuste a las especificaciones del metamodelo definido.

Objetivos Específicos

- Revisar las alternativas disponibles de Sistemas de Trazabilidad de alimentos y la información que brindan.
- Desarrollar una plataforma de auditoría de trazabilidad de alimentos que sea útil para el proceso de toma de decisiones estratégicas en la cadena de producción y comercialización de vinos.
- Fortalecer el intercambio de experiencias y avances realizados entre las instituciones participantes.
- Incrementar las actividades de investigación futura entre las universidades involucradas.
- Formulación y pruebas de reglas contra el sistema de gestión de reglas de negocio definido o diseñado para tal caso.
- Realizar ensayos y pruebas de la plataforma con los datos de dos de las bodegas relevadas. (Casos reales).

FORMACIÓN DE RECURSOS HUMANOS

En este proyecto se ha previsto la siguiente contribución a la formación de recursos humanos:

- Formación de Docentes, en el área de Auditoría de Sistemas, Calidad de Sistemas, Calidad de procesos productivos del vino y Trazabilidad, de la comunidad Universitaria de ambas Universidades.
- Contribución a Cursos de Posgrado en ambas Universidades.
- Capacitación de alumnos de grado.
- Alumnos becarios de investigación.
- Aporte a pares de investigación.
- Aporte a profesionales que se desempeñan en el medio productivo tanto en la industria del Software, respecto a la auditoría, reingeniería y calidad de Sistemas como a la industria del vino.

El proyecto integra en su equipo de investigación docentes, graduados y estudiantes investigadores de la Universidad del Aconcagua y de la Universidad Tecnológica Nacional; de las Carreras Ingeniería en Sistemas de Información, Licenciatura en Enología, Maestría en Ingeniería en Calidad, Licenciatura en Informática y Desarrollo de Software. Se trabaja con alumnos becarios. Además de investigadores y becarios de ambas Universidades, el equipo de este proyecto también está integrado por cuatro estudiantes del quinto nivel de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional quienes recibieron capacitación y nivelación para poder desarrollar el metamodelo de datos y procesos como Proyecto Final de su Carrera. Se otorgaron responsabilidades en forma colaborativa. En el estudio experimental se plasman actividades de modelado, evaluación y testing. En el desarrollo tecnológico que ejecutan tareas de análisis, diseño, desarrollo e implementación.

BIBLIOGRAFÍA

- [1] ISO 22005:2007. <https://www.iso.org/obp/ui#iso:std:iso:22005:ed-1:es:ref:3>. Último acceso 20/06/2017.
- [2] INTI, Instituto Nacional de Tecnología Industrial, <http://www.inti.gov.ar/sabercomo/sc47/inti8.php>. Último acceso 20/06/2017.
- [3] De Castro Neto, M., Rodríguez, M.B., Aguiar Pinto, P., Berger, I., "Traceability on the web – a prototype for the Portuguese beef sector", in: Proceedings of EFITA Conference, Debrecen, Hungary, 2003, pp. 607–611.358.
- [4] van Dorp, C.A "Tracking and tracing business cases: incidents, accidents and opportunities", in: Proceedings of EFITA Conference, Debrecen, Hungary, 2003, pp. 601–606
- [5] Mertens, D. M., "Research and Evaluation in Education and Psychology Integrating Diversity With Quantitative Qualitative and Mixed Methods", Tercera. SAGE Publications, 2010.
- [6] Johnson, R. B. et al., "SPECIAL ISSUE: NEW DIRECTIONS IN MIXED METHODS RESEARCH," Res. Sch., 2006.
- [7] Strauss, A.L., Corbin, J., *Bases de la investigación cualitativa: técnicas y procedimientos para desarrollar la teoría fundamentada*. Editorial Universidad de Antioquia. 2002.
- [8] Pizzuti, T., Mirabelli, G., Gómez-gonzález, G. and Sanz, M. A. "Modeling of an Agro-Food Traceability System: The Case of the Frozen Vegetables," *Proc. 2012 Int. Conf. Ind. Eng. Oper. Manag. Istanbul, Turkey*, July 3 – 6, 2012 Model., pp. 1065–1074, 2012.
- [9] Cimino, M.G.C.A., and Marcelloni, F., "Enabling traceability in the wine supply chain," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7200 LNCS, pp. 397–412, 2012.
- [10] Qasaimeh, M. and Abran, A., "An audit model for ISO 9001 traceability requirements in agile-XP environments," *J. Softw.*, vol. 8, no. 7, pp. 1556–1567, 2013.
- [11] Moguel, R.J., Vivas, O.Y., "Implementación de Trazabilidad EAN.UCC," p. 16, 2008.
- [12] "ISO 9001-Identification and Traceability." [Online]. Available: http://www.isorequirements.com/iso_9001_7.5.3_identificacion_and_traceability.html.
- [13] "Diseño de una solución de trazabilidad para la optimización del rendimiento de negocios." [Online]. Available: <https://www.ibm.com/developerworks/ssa/data/library/techarticle/dm-0902oberhofer/index.html>.
- [14] Farmer, F.Z. and Farmer, F.Z., "Good Agricultural Practices Food Safety Plan," pp. 1–21, 2014.
- [15] "Scrum." [Online]. Available: www.scrum.org.
- [16] "Taiga." [Online]. Available: <https://taiga.io>.
- [17] Object Management Group. "Business Process Model and Notation (BPMN)". OMG, (2011). <http://www.omg.org/spec/BPMN/2.0/PDF/>. Último Acceso 15/07/2017.
- [18] BPMS (Business Processes Management Suites). <http://www.gartner.com/it-glossary/bpms-business-process-management-suite/>. Último Acceso 15/07/2017.
- [19] BRMS (Business Rules Management Systems). <https://www.drools.org/>. Último acceso 15/07/2017.
- [20] Java EE. www.oracle.com/technetwork/java/javaee/overview/. Último acceso 20/04/2017.
- [21] Horakova, M., Skalska, H. "Business Intelligence and Implementation in a Small Enterprise". <http://www.sijournal.org/index.php/JSI/article/viewFile/159/114>. Journal of Systems Integration. 2013. Último acceso 15/07/2017
- [22] REST. https://es.wikipedia.org/wiki/Transferencia_de_estado_representacional. Último acceso 20/04/2017.
- [23] OpenIDconnect. <http://openid.net/connect/>. Último acceso 20/04/2017.
- [24] OAuth 2. <https://oauth.net/2/>. Último acceso 20/04/2017.
- [25] Keycloak. <http://www.keycloak.org/>. Último acceso 20/04/2017.
- [26] AngularJS. <https://angularjs.org/>. Último acceso 20/04/2017.
- [27] Moe, T.: Perspectives on traceability in food manufacture. *Food Science and Technology* 9, 211–214 (1998)
- [28] Freund, J., Rucker, B., Hitpass, B., "BPMN 2.0 Manual de Referencia y Guía Práctica". BPMCenter. Primera edición. 2011.
- [29] BonitaSoft. <http://es.bonitasoft.com/>. Último acceso 20/04/2017.
- [30] Wildfly. <http://wildfly.org/>. Último acceso 20/04/2017.
- [31] Java Persistence 2.0. Java Community Process. <http://jcp.org/en/jsr/detail?id=317> <http://docs.oracle.com/javase/6/tutorial/doc/>. Último acceso 14/7/2017.
- [32] Hibernate. <http://hibernate.org/orm/>. Último acceso 14/7/2017.
- [33] Bernabeu, R. (2010). Hefesto. Córdoba, Argentina. <https://www.businessintelligence.info/assets/hefesto-v2.pdf>. Último acceso 20/04/2017.
- [34] Pentaho. <http://community.pentaho.com/>. Último acceso 14/7/2017.

TRATAMIENTO DE EVIDENCIAS DIGITALES FORENSES EN DISPOSITIVOS MÓVILES

Liliana Figueroa, Cecilia Lara, Norma Lesca, Graciela Viaña, Adriana Binda

Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias

Exactas y Tecnologías, Universidad Nacional de Santiago del Estero

lmvfigueroa@yahoo.com.ar; {laraceciliacristina, norma.lesca}@gmail.com; gv857@hotmail.com;

abinda_arg@yahoo.com.ar

RESUMEN

En este artículo se presenta la investigación que se viene desarrollando en el Instituto de Investigaciones en Informática y Sistemas de Información de la Universidad Nacional de Santiago del Estero (UNSE) sobre obtención de evidencias digitales de dispositivos móviles. Para ello se considera como marco normativo el Nuevo Sistema Procesal Penal de la Provincia de Santiago del Estero

El proceso de adquisición de evidencias digitales debe ser legalmente aceptable, apoyándose en métodos científicos que permitan recolectar, analizar y validar las mismas, recurriendo entonces a la Informática Forense.

En este contexto, y en el marco de la investigación, se han desarrollado actividades con el propósito de definir un conjunto de lineamientos a los que se puede recurrir al momento de realizar la obtención de evidencias digitales desde dispositivos móviles. Durante esta etapa, se han estudiado las buenas prácticas tendientes a asegurar la calidad de los procesos aplicados y sus resultados.

Palabras clave:

Informática Forense, dispositivos móviles, evidencias digitales, protocolo de extracción de datos, calidad de la evidencia digital.

CONTEXTO

La presente línea de investigación se encuentra inserta en el proyecto “Computación Móvil: desarrollo de

aplicaciones y análisis forense”, que propone una continuación del trabajo en el ámbito de la computación móvil iniciada en el año 2012 [5]. Está financiada por el Consejo de Ciencia y Técnica de la UNSE. La investigación comenzó en el año 2017 y finalizará en el año 2018, y hasta la fecha se han obtenido resultados parciales en relación a los procesos que permitan garantizar la integridad de las evidencias digitales obtenidas de dispositivos móviles.

La justicia moderna necesita ampliar la mirada a la hora de obtener evidencias y pruebas digitales, que sean legalmente aceptables y que ayuden a resolver conflictos apoyándose en métodos científicos que permitan recolectar, analizar y validar pruebas digitales.

En ese ámbito, se advierte la necesidad de trabajar en el estudio y definición de un protocolo para el análisis de evidencias forenses obtenidas de dispositivos móviles. En respuesta a ello, se propone investigar: protocolos de intervenciones forenses y la gestión de evidencias digitales empleando repositorios digitales (acceso, almacenamiento, recuperación, seguridad) de las evidencias digitales.

La presente investigación se desarrolla en acuerdo y colaboración con el Gabinete de Ciencias Forenses del Ministerio Público Fiscal de Santiago del Estero, para lo cual se ha firmado un convenio que permite realizar la investigación de manera conjunta.

1. INTRODUCCIÓN

La Informática Forense es disciplina que surge a partir de sucesos que han afectado a

la sociedad globalizada e informatizada actual, en donde se observa el crecimiento una serie de delitos que están afectando diferentes áreas de la sociedad [4].

El propósito de esta disciplina es determinar los responsables de los delitos, así como también esclarecer la causa original de un ilícito o evento particular para asegurar que no se vuelva a repetir. Para ello, se encarga de recolectar pruebas digitales para fines judiciales, mediante la aplicación de técnicas de análisis y de investigación.

Haciendo una analogía con la criminalística [1], es necesario establecer un conjunto de herramientas, estrategias y acciones que ayuden a identificar hechos y evidencias relacionados con la Informática

Es por ello, que es necesario la aplicación de procedimientos estrictos y cuidadosos, desde el momento en que se realiza la recolección de la evidencia, hasta que se obtienen los resultados posteriores a la investigación [3,8]. Todo esto en el marco del Código Procesal Penal de cada provincia, ya que deben adaptar un modelo generalizado para realizar este tipo de investigaciones a las necesidades locales.

Como en cualquier investigación forense, existen una variedad de enfoques que se pueden utilizar para la recolección y análisis de información. Un aspecto clave para ello es que el procedimiento que se siga, no modifique la fuente de información de ninguna manera, o que de ser esto absolutamente necesario, el analista esté en la capacidad de justificar por qué realizó esta acción [6].

Para mantener la integridad de la evidencia y de ese modo garantizar su admisibilidad en el proceso penal se requiere de procedimientos que sigan una metodología para tal fin. En este sentido, la norma ISO/IEC 27037:2012 “Information technology - Security techniques - Guidelines for identification, collection, acquisition and preservation of digital evidence” [7] puede considerarse como marco de referencia para el tratamiento de

evidencia digital, ya que proporciona pautas para su identificación, sistematización, recolección, adquisición y preservación.

Según esta norma, la evidencia digital está gobernada por tres principios fundamentales que definen la formalidad de una investigación y son condiciones necesarias y suficientes para que se recaben, aseguren y preserven elementos probatorios sobre medios digitales. Éstos son:

- *Relevancia*: condición técnicamente jurídica que habla sobre los elementos pertinentes a la situación que se analiza, con el fin de probar o rechazar una hipótesis sobre los hechos.
- *Confiabilidad*: acción orientada a saber si la evidencia que se extrae u obtiene es lo que debe ser. Se busca validar la repetibilidad y la auditabilidad del proceso aplicado para obtener la evidencia digital.
- *Suficiencia*: condición por la que con las evidencias recolectadas y analizadas hay suficientes elementos para sustentar los hallazgos y verificar las afirmaciones sobre la situación investigada.

En la actualidad, el empleo de dispositivos móviles se ha incrementado notablemente, principalmente por su facilidad de uso y la propiedad de mantener en contacto permanente a sus usuarios. A partir de esto, se ha generado un cambio significativo en la forma en que las personas se comunican, pero también por su proliferación, se ha incrementado su uso en actividades de orden delictivo.

A diferencia de la Informática Forense clásica, el análisis forense sobre dispositivos móviles, es un campo relativamente nuevo y los procedimientos y normas para su análisis aún se encuentran en desarrollo [11]. Un análisis forense que se lleve a cabo sobre un dispositivo móvil, puede ser admitido o no en un juicio dependiendo de lo que considere el juez y la formalidad con que se desarrolle el procedimiento de recolección,

control, análisis y presentación de las evidencias.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Considerando la amplitud de los aspectos relacionados con la computación móvil, la línea de investigación se refiere a:

Informática Forense: protocolo y gestión de evidencias digitales obtenidas de dispositivos móviles.

A partir de ella, se proponen dos líneas de investigación derivadas, consideradas desde el ámbito de la justicia penal de Santiago del Estero:

- Protocolo de actuación para la extracción de evidencias digitales de dispositivos móviles.
- Modelo de datos para la gestión de las evidencias.

El objetivo general de la investigación propuesta relacionada a esta línea de investigación es:

- *Contribuir al progreso del campo de la Computación Móvil mediante el análisis forense de dispositivos móviles.*

Los objetivos específicos que permitirán alcanzar el objetivo general son:

- *Definir un protocolo de actuación en la extracción de evidencias digitales de dispositivos móviles en el marco del nuevo Código Procesal Penal de la provincia de Santiago del Estero.*
- *Diseñar un repositorio de evidencias digitales extraídas de dispositivos móviles en el marco del proceso penal mencionado.*

Se trata de una investigación descriptiva-cualitativa, dado que si bien se definió una hipótesis que relaciona variables, la misma no alcanzará a ser corroborada en el plazo de dos años que dura esta investigación.

La hipótesis planteada es la siguiente:

El uso de un protocolo preestablecido de Informática Forense para móviles y de un repositorio especializado, optimiza la gestión de evidencias digitales extraídas de los dispositivos móviles.

Como puede observarse en la misma, la variable a estudiar es la “optimización de la gestión de evidencias criminales obtenidas de dispositivos móviles”, la cual en futuras investigaciones podrá ser evaluada a través de indicadores cuantitativos que se pueden aplicar a casos de prueba especialmente diseñados.

4. RESULTADOS OBTENIDOS

A la fecha se obtuvieron los siguientes resultados parciales:

1. *Relevamiento y análisis comparativos de protocolos y guías nacionales e internacionales para el tratamiento de la evidencia digital.* Tal como se destaca en [9], se han analizado las fases definidas en el Proceso Unificado de Recuperación de la Información PURI, que brinda una visión detallada y abarcadora de todo lo concerniente a la labor relacionada a la adquisición de evidencias digitales. Allí también, se consideran las normas ISO/IEC 27000, las cuales brindan una serie de definiciones relacionadas a la evidencia digital y establecen los principios fundamentales que definen la formalidad de una investigación: relevancia, confiabilidad y suficiencia.

2. *Análisis jurídico-legal sobre tratamiento de evidencias digitales.* Se ha analizado normativa y jurisprudencia nacional e internacional relacionada con el tratamiento de la evidencia digital sobre dispositivos móviles. Se han estudiado comparativamente diferentes códigos de procedimiento, analizando sus regulaciones sobre el proceso de obtención de la evidencia digital.

3. *Relevamiento de las herramientas de hardware y software que se usan para obtener evidencias de móviles.*

4. *Lineamientos para el tratamiento de evidencias digitales forenses sobre*

dispositivos móviles: se han identificado fases, a saber: recepción del requerimiento judicial, recolección y preservación de la evidencia digital, identificación del requerimiento judicial, adquisición de la copia forense, extracción y análisis de la evidencia digital, preparación del informe o dictamen pericial y la remisión de los elementos peritados y entrega del dictamen. Estas fases intentan garantizar el cumplimiento de las buenas prácticas que aseguren la calidad de los procesos aplicados y sus resultados, considerando los principios fundamentales de relevancia, confiabilidad y suficiencia establecidos en la norma ISO/IEC 27037[7].

A futuro, se pretende aplicar una lista de verificación con preguntas relacionadas con cada uno de los principios que establece la norma, debido a que el mencionado estándar sólo se describen los principios, pero no se especifican líneas de acción para llevarlos a cabo, a partir de las cuales se puedan derivar los mecanismos de validación asociados [2]

Al finalizar la investigación se espera contar con nuevo conocimiento científico-tecnológico, plasmado en un protocolo para la recolección y tratamiento de evidencias digitales criminales extraídas de dispositivos móviles, acompañado de un modelo para la gestión óptima de dichas evidencias en el ámbito del Poder Judicial y del Ministerio Público Fiscal de la Provincia de Santiago del Estero, y de acuerdo a lo establecido en el nuevo Código Procesal Penal de la provincia [10].

Se considera que la obtención del mencionado protocolo traerá un beneficio muy importante para la justicia santiagueña, dado que actualmente no existe un procedimiento claro y definido. Permitiría mejorar la calidad de las evidencias digitales y ayudará en la labor de los fiscales de la provincia.

5. FORMACIÓN DE RECURSOS HUMANOS

La Directora y Codirectora del proyecto pertenecen al Departamento de Informática de la UNSE. Los asesores pertenecen a LIDI-FI-UNLP y FCE-UNSa.

Los investigadores de esta línea específica de Informática Forense constituyen un equipo interdisciplinario conformado por cinco docentes de la UNSE y de la Universidad Nacional de Salta (UNSa), con profesión en Informática y Derecho. Estos poseen distintas categorías de investigación y algunos desempeñan sus actividades profesionales en: Poder Judicial de la Provincia de Santiago del Estero, Gabinete de Ciencias Forenses del Ministerio Público Fiscal de Santiago del Estero y Poder Judicial de la Provincia de Salta.

El equipo de investigación se encuentra asistiendo y asesorando a alumnos de grado y posgrado de UNSE y UNSa que realizan sus trabajos de finalización de carrera en temáticas relacionadas con esta línea de investigación. Estos se encuentran en la etapa de propuesta inicial.

6. REFERENCIAS

1. CANO, J. (2006). Introducción a la informática forense: Una disciplina técnico-legal. Revista Sistemas, Asociación Colombiana de Ingenieros de Sistemas (ACIS). Vol.96, pp. 64-73. http://52.0.140.184/typo43/fileadmin/Revista_96/dos.pdf
2. CANO, J. (2013) IT-Insecurity. Disponible en <http://insecurityit.blogspot.com.ar/2013/09/reflexiones-sobre-la-norma-isoiec.html>
3. CASTILLO, C., ROMERO, A., CANO, J. (2008). Análisis Forense Orientado a Incidentes en Teléfonos Celulares GSM: Una Guía Metodológica. Conf. XXXIV Conferencia Latinoamericana de Informática, Centro Latinoamericano de Estudios en Informática (CLEI). <http://www.clei2008.org.ar>.

4. DARAHOGE, M. (2011). Manual de Informática Forense. Buenos Aires. Errepar.
5. HERRERA, S., NAJAR RUIZ P., ROCABADO S., FENNEMA, C., CIANFERONI, M. (2013). Optimización de la calidad de los sistemas móviles. http://sedici.unlp.edu.ar/bitstream/handle/10915/27200/Optimizaci%C3%B3n_de_la_calidad_de_los_sistemas_m%C3%B3viles.pdf?sequence=1
6. HOOG, A. (2009). iPhone Forensics: Annual Report on iPhone Forensic Industry. Chicago Electronic Discovery.
7. ISO/IEC 27037:2012(en) Information technology— Security techniques— Guidelines for identification, collection, acquisition and preservation of digital evidence. <https://www.iso.org/obp/ui/#iso:std:iso-iec:27037:ed-1:v1:en>
8. LEIGLAND, R. (2004). A Formalization of Digital Forensics. International Journal of Digital Evidence. University of Idaho. Volume 3, Issue 2. <http://www.utica.edu/academic/institutes/ecii/publications/articles/A0B8472C-D1D2-8F98-8F7597844CF74DF8.pdf>.
9. LESCA, N., LARA, C., FIGUEROA, L., VIAÑA, G. (2017). Gestión de evidencia digital en dispositivos móviles. Jornadas Argentinas de Informática - Simposio Argentino de Informática y Derecho. ISSN: 2451-7526
10. LEY 6.941. (2009). Código Procesal Penal de Santiago del Estero. Disponible en <http://www.jussantiago.gov.ar/jusnueva/Normativa/Ley6941.pdf>
11. VARSALONE, J., KUBASIAK, R. (2009). Mac Os X, iPod and iPhone Forensic Analysis DVD Toolkit. Syngress Publishing, Inc, pp. 355-475.

Juegos serios y aplicaciones interactivas usando realidad aumentada y realidad virtual

Javier Diaz, Laura Fava, Claudia Banchoff, Alejandra Schiavoni, Sofía Martin

LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.
Facultad de Informática. Universidad Nacional de La Plata
Calle 50 esq. 120, 2do Piso. Tel: +54 221 4223528
{jdiaz, lfava, cbanchoff, ales}@info.unlp.edu.ar, smartin@linti.unlp.edu.ar

Resumen

Este artículo describe la línea de investigación, desarrollo e innovación "Juegos Serios y Aplicaciones Interactivas usando Realidad Aumentada y Realidad Virtual" que se viene desarrollando en el LINTI, Laboratorio de Investigación y Nuevas Tecnologías Informáticas de la UNLP (Argentina).

Dentro de esta línea se abordan aspectos relacionados al desarrollo de aplicaciones principalmente de uso educativo que utilizan Realidad Aumentada (RA) y Realidad Virtual (RV), tales como recursos educativos, materiales didácticos, objetos de aprendizaje y videojuegos educativos, y también aquellas destinadas a otras áreas como turismo y edificios históricos interactivos, entre otras. En esta etapa se hará enfoque en el análisis y evaluación de entornos de desarrollo, analizando aspectos de usabilidad, flexibilidad y adecuación, especialmente para aplicaciones de gamificación. A partir de esta evaluación, se seleccionará una herramienta para desarrollar prototipos sobre diversas aplicaciones interactivas, tales como un juego de tablero para enseñar historia destinado a escuelas secundarias y una aplicación para ser utilizada en el Museo de Ciencias

Naturales de la Universidad Nacional de La Plata.

Palabras clave: juegos serios, aprendizaje basado en juegos, realidad aumentada, realidad virtual, gamificación.

Contexto

En el LINTI se está trabajando con la temática videojuegos y juegos serios desde el año 2009, con diferentes iniciativas, todas ellas vinculadas a problemáticas reales y que representan un elemento tecnológico innovador en el aula de la escuela y en ámbitos educativos no formales. Se han implementado juegos serios de diferentes géneros en el marco del desarrollo de tesinas de grado, entre los que se destacan eQuino (Díaz J., Queiruga C., Arce T., 2013) y CarToon (Farinella M. p & Ive, S., 2014), dos videojuegos inmersivos, y RAÍCES (Díaz, J., Fava, L., Nomdedeu, L., 2013) un videojuego de plataforma. También se desarrollaron videojuegos educativos en el marco de algunas asignaturas de grado, como Zedpy¹ desarrollado en la materia Seminario de Lenguajes (opción Python). En el caso de RAICES² se cuenta con

¹ <https://github.com/Robots-Linti/Zedpy>

² <http://raiceseljuego.com.ar>

resultados de uso en experiencias muy alentadoras (Fava, L., Banchoff, C., Nomdedeu, L., Martín, S., 2017). Zedpy y otros trabajos similares realizados en el Seminario de Lenguajes, fueron presentados en congresos de videojuegos y educación (Banchoff C, Vázquez M., Harari V. & Martín S, 2016).

Algunos de los docentes que trabajan en esta temática participan en proyectos de extensión con vinculación a instituciones educativas. Esto permite contar con un campo de aplicación real para probar las aplicaciones desarrolladas y analizar el impacto con estudiantes y docentes.

En el caso de la aplicación diseñada para el Museo de Ciencias Naturales, la misma se enmarca en un proyecto de dicha institución en el cual se está renovando una sala de paleontología. Asimismo, algunas de las ideas y desarrollos pueden aplicarse a otros museos de la red de museos de la universidad.

La línea de investigación "Juegos Serios y Aplicaciones Interactivas usando Realidad Aumentada y Realidad Virtual" está enmarcada en el proyecto "Internet del Futuro: Ciudades Digitales Inclusivas, Innovadoras y Sustentables, IoT, Ciberseguridad, Espacios de Aprendizaje del Futuro" del Programa Nacional de Incentivos a docentes-investigadores, que se desarrolla en el LINTI.

Introducción

Es sabido que los métodos tradicionales de enseñanza como los libros o la disertación oral, muchas veces fallan al momento de motivar a sus destinatarios y capturar su atención. Los videojuegos y las aplicaciones interactivas poseen una

motivación intrínseca muy interesante para ser usadas en múltiples contextos de enseñanza-aprendizaje como complemento motivacional. Si además se incorporan aspectos de realidad aumentada (RA) y/o realidad virtual (RV), podrían convertirse en instrumentos muy poderosos para ser usados en múltiples contextos.

Los conceptos de RA y RV están muy relacionados, podemos decir que la realidad es todo aquello que se puede percibir por medio de los sentidos (directa o indirectamente), lo virtual es algo que percibimos pero que no se corresponde con la realidad en ese espacio y tiempo (espejismo, grabación virtual, etc.).

En 1994, Paul Milgram y Fumio Kishino (Milgram P. & Kishino F., 1994) presentaron el concepto del continuo de la virtualidad en el Reality-Virtuality Continuum, como se muestra en la Fig. 1.



Fig 1: RA y RV en base al Reality-Virtuality Continuum

La RA combina la realidad con información sintética, imágenes 3D, videos, texto y también sonidos y sensaciones táctiles (*haptics*) en tiempo real teniendo en cuenta la posición y dirección de la visión del usuario. La RA integra elementos reales y virtuales, pero considerándola más cercana al mundo real, en esta combinación de ambos

mundos, no se tiene la intención de suprimir el mundo real en el que nos desenvolvemos, sino por el contrario, aumentarlo.

La RV está en el otro extremo, es un mundo generado por computadora con el cual un usuario puede interactuar, la interacción puede variar desde simplemente mirar hasta modificar interactivamente el mundo.

La RA se aplica desde hace varios años. En (Chen P., Liu X., Cheng W., Huang R , 2017), se analizaron 55 estudios sobre el uso de RA en educación entre los años 2011 y 2016. En dicho análisis se detallan los campos más importantes en que se ha utilizado efectivamente la RA, como ser: ciencia (experimentos de laboratorio), enseñanza de matemática y geometría, geografía y ecología, enseñanza de idiomas, apreciación de arte visual y pinturas, ingeniería y salud. En este trabajo se concluyó que la aplicación de técnicas de RA en educación conducen a un mejor rendimiento en el proceso de aprendizaje promoviendo un mayor compromiso de los estudiantes y una mayor motivación en el proceso. En este mismo análisis, se detectó que las técnicas de RA basada en el reconocimiento de imágenes son más utilizadas que las basadas en ubicación, y que el aula es el entorno más utilizado para su aplicación. Se han realizados algunas experiencias donde se implementó RA en museos. En (Kuo-En Chang, Chia-Tzu Chang, 2013) se realizó una experiencia con un grupo de 130 personas utilizando diferentes métodos de guías para acompañar la visita a una misma sala, como ser guías con RA, audio guías y sin guías. De los datos obtenidos, una de las conclusiones más destacadas es que la utilización de RA

permitió que los usuarios incorporaran más información sobre los objetos observados sin perder la conexión entre lo virtual y el objeto real.

Si bien las mismas ventajas podrían aplicarse a la RV, su aplicación efectiva es más compleja debido al costo de los recursos de hardware requeridos. Una aplicación con RA puede ser usada desde un teléfono inteligente y/o tableta de uso común.

Líneas de Investigación, Desarrollo e Innovación

Los ejes en los que se está investigando están relacionados con el desarrollo de videojuegos y aplicaciones interactivas usando RA y RV para ámbitos educativos y de turismo.

A continuación se detallan los ejes de investigación:

- Análisis de herramientas y librerías existentes para el desarrollo de aplicaciones de RA y RV, con especial foco en aquellas que permitan generar aplicaciones multiplataforma.
- Capacitación en entornos para desarrollo de aplicaciones usando RA y RV como ser Unity3D y Unreal.
- Definición y desarrollo de prototipos de aplicaciones y videojuegos educativos.
- Aplicación de los prototipos desarrollados en contextos reales, a través del uso en instituciones de distintos ámbitos con el objetivo de analizar su impacto.

Resultados y Objetivos

El objetivo general de esa línea de trabajo es aplicar técnicas de RA y RV en aplicaciones y juegos serios, con especial foco en el ámbito educativo y turismo.

Para cumplir con el objetivo general, se proponen los siguientes objetivos específicos:

- Realizar un análisis detallado de las tecnologías disponibles para el desarrollo de aplicaciones con RA y RV, verificando su adecuación a los distintos tipos de aplicaciones.
- Analizar los distintos contextos de uso de las aplicaciones, los cuales direccionarán las tecnologías a utilizar.
- Desarrollar prototipos de juegos serios y aplicaciones interactivas que puedan utilizarse en el ámbito de la escuela.
- Construir casos de prueba de los diferentes prototipos con el fin de aplicarlos en entornos educativos reales, para comprobar su adecuación y usabilidad.
- Promover esta temática en el marco del desarrollo de las tesis de grado.

Además de los videojuegos mencionados en el contexto de este artículo, se está trabajando en el desarrollo de dos aplicaciones interactivas usando RA. Una de ellas con personas ilustres de la UNLP y la otra para la enseñanza de historia en la escuela secundaria. En la Fig. 2 se muestran

bocetos de cada uno de los ejemplos mencionados.

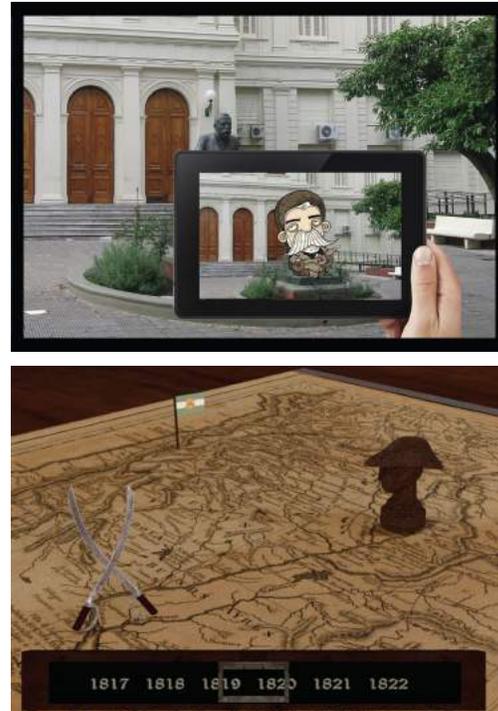


Fig. 2: Diseño de aplicaciones con RA

En estos casos, se prevé finalizar ambos desarrollos y realizar una prueba de campo en algunas de las instituciones educativas con las que se trabaja.

Para llevar a cabo los objetivos planteados, se realizaron diversas actividades a fin de formar un equipo de trabajo especializado. Durante el mes de noviembre de 2017 un profesor de la Universidad de Ciencias Aplicadas NHTV-Breda, Holanda, dictó dos talleres de desarrollo de videojuegos con RA y RV en el cual participaron docentes y alumnos. En la semana del 11 de diciembre del mismo año, un profesor de

la Universidad Politécnica de Valencia dictó un curso intensivo sobre desarrollo de videojuegos con RA y RV usando Unity3D. Estos talleres permitieron formar un equipo de trabajo que no sólo permitirá diseñar e implementar las aplicaciones planificadas, sino que también contribuirá en la capacitación en esta temática. En este sentido, se prevé el dictado de un curso en la Universidad de Salta y otro en el ámbito de la Facultad de Informática, destinado a todos aquellos que, por diversas cuestiones no pudieron concurrir a los mismos en el año 2017.

Formación de Recursos Humanos

El equipo de trabajo de la línea de I+D+i presentada en esta artículo está formado por docentes investigadores categorizados del LINTI y estudiantes de la Facultad de Informática. Asimismo en algunos de los proyectos se trabajará con un equipo interdisciplinario de profesionales del Museo de Ciencias Naturales de la UNLP. A través de la generación permanente de conocimiento por medio de líneas de investigación y desarrollo de aplicaciones vinculadas al sector productivo y su aplicación en el ámbito social, el LINTI promueve el uso innovador de las tecnologías informáticas en la región.

En relación a las tesinas de grado vinculadas con esta línea de investigación, se está desarrollando una aplicación interactiva para el Museo de Ciencias Naturales y un desarrollo para el Hospital Garrahan.

Referencias

Banchoff C, Harari V. &, Martin S. (2015) Desarrollo de videojuegos educativos para los primeros años de escuela. Una forma de integrar docencia, extensión e investigación en la Facultad

de Informática. III Congreso Internacional de Videojuegos y Educación 2015 (CIVE15). Disponible en <https://goo.gl/Wo5w3W>.

Chen P., Liu X., Cheng W., Huang R. (2017) A review of using Augmented Reality in Education from 2011 to 2016. In: Popescu E. et al. (eds) Innovations in Smart Learning. Lecture Notes in Educational Technology. Springer, Singapore.

Díaz, J., Fava, L, Nomdedeu, L (2013), Design of a Social Serious Game to revalue the Argentinian aboriginal cultures, International Conference on Collaboration Technologies and Systems, CTS 2013. ISBN: 978-1-4673-6402-7.

Díaz J., Queiruga C., Arce T. (2013). e-Quino: an Interactive Videogame to Complement Equine Therapy (2013). Proceedings of the 8th Immersive Education Summit (iED 2013), Boston MA. Immersive Education Initiative. ISSN 2325-4041.

Fava, L., Banchoff, C., Nomdedeu, L., Martín, S. (2017) Las culturas originarias a través de Raíces Un videojuego como recurso didáctico-cognitivo. LACLO 2017. ISBN: 978-1-5386-2376-3.

Farinella M. p & Ive, S. (2014). CarTON: manejando en la ciudad. Disponible en <http://sedici.unlp.edu.ar/handle/10915/475>.

Banchoff C, Vazquez M., Harari V. &, Martin S. (2016).(Banchoff C, Harari V.

&, Martin S, (2016)
<https://cive2016creas2i.files.wordpress.com/2017/06/delaideaalapantalla-disencc83o-de-juegos-serios.pdf>.

Milgram P. & Kishino F. (1994). A Taxonomy of Mixed Reality Visual Displays. IEICE Transactions on Information Systems, E77-D (12), diciembre 1994.

Kuo-En Chang, Chia-Tzu Chang, Huei-Tse Hou, Yao-Ting Sung, Huei-Lin Chao, Cheng-Ming Lee. (2013). Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum.

ROBÓTICA BASADA EN INTERNET DE LAS COSAS

Azurra, D., Santos D., Fernández, G., Fernández S.

Dpto. Desarrollo Productivo y Tecnológico
Universidad Nacional de Lanús.
29 de Septiembre 3901 (1826) Remedios de Escalada, Lanús
Buenos Aires, Argentina.

Resumen

Este proyecto se inscribe en una línea de investigación que busca desarrollar y sistematizar el cuerpo de conocimiento de la administración y control de robots utilizando plataformas de IoT, con focalización en su transferencia a la Industria, particularmente al sector PyME.

En este trabajo se presentan supuestos que guían al proyecto, así como las actividades, objetivos y resultados esperados con relación a esta investigación.

Palabras clave: sistemas embebidos - robótica - Internet de las Cosas - telecomunicaciones - sistemas industriales

Contexto

Este proyecto de investigación integra la línea de trabajo en aplicaciones de sistemas industriales, robótica y telecomunicaciones en el marco de la carrera de la Carrera de Licenciatura en Sistemas de la Universidad Nacional de Lanús.

Introducción

Según [1], en los últimos años la robótica ha tenido aplicaciones en campos tan diversos y críticos como la medicina, la exploración planetaria y submarina, automatización de procesos industriales, seguridad, entretenimiento, entre otros. Sin embargo, es en la educación donde ha dado uno de los aportes de mayor impacto, donde los robots al

integrarse al grupo de estudiantes y tutores, propician el aprendizaje y el fortalecimiento de habilidades cognitivas.

Por su parte, según la definición de la Internet of Things Community del Instituto de Ingeniería Eléctrica y Electrónica (IEEE), Internet de las Cosas (IoT, por sus siglas en inglés) es un concepto de computación en el que todas las cosas, incluyendo todo objeto físico, pueden ser conectadas, lo que los hace inteligentes, programables y capaces de interactuar con los seres humanos.[2]

Avanzando con el concepto, en [3] se destaca que la IoT involucra sensores, circuitos, sistemas embebidos, comunicaciones, interfaces inteligentes, gestión de energía, gestión de datos, fusión de datos, gestión de conocimiento, sistemas en tiempo real, procesamiento distribuido, diseño de sistemas y técnicas sofisticadas de software que se relacionan con la llamada Big Data. También prevé que su desarrollo podría cambiar profundamente desde los procesos productivos hasta la salud-electrónica, ya que ofrece nuevas formas para el cuidado de humanos y el tratamiento de dolencias.

Un análisis desde el punto de vista económico es realizado en [4], previendo una inversión de 17.000 millones de dólares de aquí a 2020, definiéndolo como un mercado en expansión, donde cada vez son más las empresas que valoran la conectividad entre dispositivos y con la Red.

En [5] se definen a las plataformas de IoT como la base para que los dispositivos estén interconectados y se genere un ecosistema propio. Detalla que constituyen el software al que se conectan los dispositivos de hardware,

brindando comunicación y puntos de acceso para el desarrollo de aplicaciones.

Líneas de investigación y desarrollo

Este proyecto se inscribe en una línea de investigación que busca desarrollar y sistematizar el cuerpo de conocimiento de la administración y control de robots utilizando plataformas de IoT, con focalización en su transferencia a la Industria, particularmente al sector PyME.

Entre los supuestos que guían al proyecto se encuentran:

- I. Existen plataformas de IoT robustas, fiables y de acceso libre o bajo costo que podrían ser utilizadas en la administración y control de robots.
- II. Es posible adaptar los mecanismos de administración y control de robots para que interactúen con estas plataformas.
- III. Es factible desarrollar un modelo de transferencia a la Industria, particularmente al sector PyME.

Resultados y Objetivos

El objetivo general de este proyecto es el estudio de plataformas de IoT utilizables en robótica, buscando: [a] caracterizar plataformas de IoT disponibles; y [b] identificar e implementar las modificaciones necesarias para adaptar los mecanismos de administración y control de robots a esta tecnología.

Este objetivo general se desarrollará mediante los siguientes objetivos específicos:

- I. Relevar diferentes plataformas de IoT disponibles, identificando características específicas y diferenciales, y determinando la aplicabilidad de las mismas en la administración y control de robots.

II. Seleccionar plataformas de IoT, e implementar en éstas algoritmos de administración y control de robots.

III. Identificar las modificaciones necesarias para adaptar los mecanismos de administración y control de robots para que interactúen con plataformas de IoT.

IV. Desarrollar prototipos de software y hardware de mecanismos de administración y control de robots para que interactúen con las plataformas de IoT seleccionadas, y así poner a prueba a las mismas.

V. Desarrollar un modelo de aplicación y uso de esta tecnología en la industria, particularmente en el sector PyME.

Formación de Recursos Humanos

El grupo de trabajo está formado por cuatro docentes-investigadores de la carrera de Licenciatura en Sistemas de la Universidad Nacional de Lanús.

Asimismo, dos alumnos avanzados están desarrollando su trabajo de fin de carrera en el marco del proyecto.

Referencias

- [1] J. González E.; B. Jovani A. Jiménez ; (2009); “LA ROBÓTICA COMO HERRAMIENTA PARA LA EDUCACIÓN EN CIENCIAS E INGENIERÍA”; Revista Iberoamericana de Informática Educativa, 10: 31-36; ISSN: 1699-4574
- [2] IEEE Internet of Things Community ; <https://www.ieee.org/>
- [3] Vincenzo Piuri, Roberto Minerva, Construyendo la Internet de las Cosas, Julio 2015; <https://www.computer.org/web/computingnow/archive/building-the-internet-of-things-july-2015-spanish-version>

[4] El Internet de las Cosas de código abierto: plataformas y aplicaciones para desarrolladores, Agosto 2015;
<https://bbvaopen4u.com/es/actualidad/el-internet-de-las-cosas-de-codigo-abierto>

[5] ¿Qué es una plataforma IoT?, 2016 ;
<https://secmotic.com/blog/plataforma-iot/>

PROPUESTA DE RECOLECCIÓN DE DATOS PARA AFORO VEHICULAR

Julio MONETTI, Micaela CONTRERAS, Martín NAVARRO & Gerardo SEVILLANO
Ingeniería en Sistemas de Información/FR.Mendoza -UTN
{micaela.contreras124, navarromartin762, gero.sevillano, julio.monetti}@gmail.com

RESUMEN

El estudio de la circulación vehicular en una ciudad permite determinar la conveniencia en la construcción de nuevas vías, mejorar las existentes, asignar o modificar rutas en el transporte público, determinar controles de tránsito y transporte en sitios y periodos específicos, entre otras. Todo esto tiene como objetivo principal mejorar la calidad de transporte en zonas críticas de la ciudad. Para ello, el presente estudio pretende avanzar con tareas de recolección automática de datos, modelado y simulación para satisfacer este objetivo. Para ello se aborda la construcción de un modelo de datos robusto que permita mantener la información actualizada en todo momento.

Palabras clave: tráfico vehicular, microsimulación.

CONTEXTO

El presente trabajo se encuadra en las actividades de investigación y desarrollo del grupo GICAR (Universidad UTN-FRMendoza), el cual proyecta actividades de simulación de tráfico vehicular para la descripción de escenarios particulares. Actualmente el grupo se encuentra organizando el conjunto de datos necesarios para la calibración el modelo. La adquisición de datos conforman dos ejes principales: 1) La descripción de la geometría del área a simular dada por la georeferenciación de calles, cruces, etc; y 2) La descripción del movimiento de vehículos (agentes), a lo largo de las diferentes zonas y subzonas de estudio.

1. INTRODUCCION

El trabajo se centra en establecer una relación entre la geometría del área de estudio y los datos sobre la circulación de los agentes (ver figura 1). Es menester obtener información acerca del movimiento de los agentes a lo largo de las diferentes zonas de la ciudad, y poder describir fielmente diferentes variables que manifiesten situaciones bajo estudio, como por ejemplo zonas de congestión vehicular, zonas atractoras y generadoras de tráfico, etc.

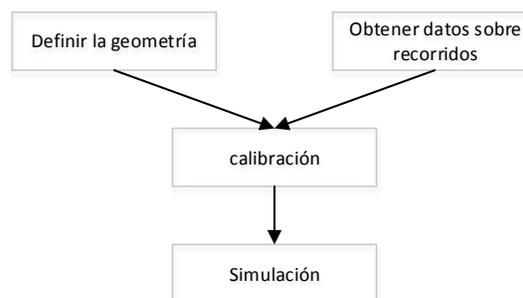


Figura 1. Resumen del proceso de adquisición y modelado de datos para la simulación de la dinámica vehicular.

El aporte de la investigación pretende establecer un modelo de datos que permita un adecuado análisis y procesamiento para facilitar el mantenimiento de la información a lo largo del tiempo.

En la sección 2 se comenta la línea de investigación actual y los resultados parciales obtenidos por el grupo a través de la representación del área de estudio [1][2]. En la sección 3 se mencionan las actividades proyectadas para la recolección de datos provenientes del aforo vehicular. Finalmente se resumen resultados y la planificación del trabajo futuro.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Se aborda en primer lugar la determinación del área de estudio a través del modelado de datos. El área de estudio esta descrita a través de un grafo dirigido que determina las posibilidades de circulación de los agentes.

Para el modelado del área se prevé la adquisición y procesamiento de datos referentes a cruces y tramos (secciones de calles de esquina a esquina). Estos datos georeferenciados representan uno de las principales entradas al sistema de simulación propuesto.

En una primera instancia, los autores no contaron con datos específicos suficientes para describir el escenario a ser simulado, puesto que, por ejemplo, la adquisición directa de esta información a partir de los mapas digitales alojados en dispositivos GPS comunes en el mercado no resultó factible. Por esta razón se han propuesto diferentes metodologías para la obtención de dicha información. En [1] se utilizó información proveniente de una traza GPS proveniente de un vehículo en circulación para delinear el paso entre esquinas contiguas, proceso a través del cual se concluye con la direccionalidad de cada tramo. Esta metodología brinda resultados aceptables, pudiendo enlistar pares ordenados (*esquina origen, esquina destino*) para la conformación del grafo. Sin embargo, el procedimiento resulta poco práctico, ya que es necesario recorrer la totalidad del área de estudio para obtener todos los pares ordenados. Es en base a esto que se aborda la consulta automatizada de una base de datos abierta y disponible en Internet: OSM (*Open Street Maps*) [2], con lo cual se torna necesario definir un sistema capaz de reconocer el formato de los datos obtenidos, y la posterior explotación de los mismos. Luego esta información ya procesada es transferida a una base de datos relacional para un acceso más adecuado.

El procesamiento principal radica en la generación de un modelo de datos orientado a objetos a partir del modelo jerárquico provisto por OSM (archivo XML). Se realiza un preprocesamiento sobre el archivo XML para

obtener una lista de nodos (cruces) y aristas (tramos) ordenados adecuadamente, para lo cual se utiliza una API provista por el proyecto OSM, que es enlazada con el conjunto de algoritmos desarrollados por el grupo. A fin de complementar la funcionalidad provista por esta API, se desarrollan algoritmos que posibilitan, entre otras cosas, determinar distancias entre puntos consecutivos. Esto viabiliza asegurar la proximidad de dos nodos contiguos, descartar puntos discordantes, nodos repetidos en el conjunto de datos, etc. En este momento el grupo encuentra que es necesario un análisis previo de los datos y la confección de algoritmos de refinamiento de los datos.

Esta información se complementa con la zonificación de las diferentes subáreas, brindando a cada una de ellas un código de zona. La división política de las zonas, así también como los límites naturales - vías ferroviarias, ríos, carreteras principales, etc. - son considerados delimitadores de zonas. (En futuros trabajos se considerará el tamaño de la zona de acuerdo a la densidad poblacional.). Otro criterio válido es la determinación de zonas y su tamaño en función del tiempo que tarda un agente en transitar a través de ella.

Una vez descrita las posibilidades de circulación a través del grafo, es necesario el modelado de información sobre las características dinámicas del modelo: la circulación de los agentes.

3. AFORO VEHICULAR

En ingeniería de tránsito existe una amplia gama de metodologías para la realización de encuestas y conteo [3], que asisten al analista en el entendimiento de la dinámica dada en el transporte. En trabajos de campo previos se observa que muchos de los estudios son intervencionistas, esto es: el conteo afecta directamente el flujo vehicular, situación que es analizada en la confección de las encuestas origen-destino propuestas.

Se consideran en primer lugar los fundamentos teóricos del tráfico, y las características de circulación. El campo de aplicación de los

datos obtenidos en el aforo es sumamente extenso y aplicable por diferentes organismos dedicados al estudio de la dinámica vehicular. Es por ello que sea cual fuere la metodología de recolección adoptada se requieren delinear protocolos que aseguren la confiabilidad de los datos.

Para este estudio, el principal objetivo del conteo es la determinación de volúmenes de tránsito en la red vehicular. Luego, el conocimiento del volumen y tipo de vehículos que circulan en la red permite determinar entre otras variables, el grado de congestión por zona, condiciones de espera por tramo, y la evolución del transporte en general.

Como se mencionó anteriormente, existen diferentes modalidades para la recolección de datos, dependiendo de los objetivos del estudio cual se adopta. Se abordan especialmente en este estudio dos tipos de recolección de datos:

1. Encuestas origen-destino.

Corresponde a la interpelación del agente sobre el recorrido a realizar.

2. Conteo *in-situ*.

Corresponde al conteo vehicular clásico, donde el recolector de datos se ubica en una intersección, y a través de planillas previamente confeccionadas asienta la cuenta de vehículos.

2.1 Encuestas Origen-Destino.

En primer lugar, y teniendo en cuenta el análisis de encuestas realizadas con anterioridad y provenientes de la literatura, el grupo realiza un diseño que resulta de utilidad para el estudio proyectado. En la mayoría de los casos resulta imposible llevar a cabo una encuesta de la totalidad de agentes circulando, por lo que se torna necesario un correcto procesamiento estadístico de las muestras. Por otro lado se presumen ciertos los datos aportados por el agente, confiando en sus respuestas, lo que conlleva a realizar las mismas de una manera simple e imparcial.

De acuerdo a las características de recolección de datos, donde posiblemente la encuesta deba

ser respondida por agentes que se encuentra circulando, se propone un modelo de encuesta desdoblada en dos partes, compuesta por una *información primaria*, y una *información secundaria*. La primera corresponde a la indagación sobre la dirección origen y dirección destino (compuesta por los datos *calle, número, localidad*), y el día, hora y periodicidad del viaje. Esta información permite configurar directamente algunas de las variables más importantes para la simulación: *centros generadores y atractores de viajes*. La figura 2 muestra en forma gráfica dos centros atractores de flujo vehicular.



Figura 2. Los círculos rojos muestran centros atractores de viajes.

Los puntos intermedios de circulación entre origen y destino pueden ser obtenidos a través del procesamiento del grafo dirigido descrito en la sección 2.

La recolección y procesamiento de los datos secundarios no es crítica, pero permite en el caso de ser obtenida ajustar mejor la información procesada. En la segunda parte de la encuesta se incorporan preguntas que requieren que el entrevistado mencione puntos intermedios y/o alternativos de circulación para el viaje en cuestión.

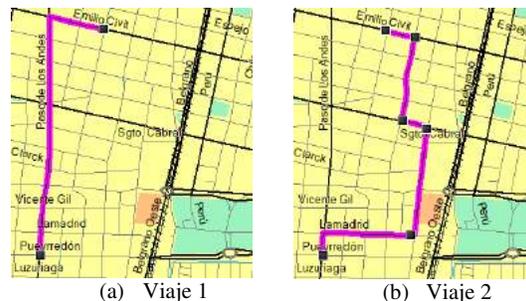


Figura 3. Caminos alternativos entre origen y destino

Conforman también el conjunto de datos secundarios

- el número y tipo de viajes,
- el propósito del viaje,
- el costo del viaje ,
- el tiempo esperado de viaje,
- el número de pasajeros movilizados,
- etc.

Resulta de especial importancia definir el lugar y el método para la recolección de datos a través de la encuesta. Como se comentó en párrafos anteriores, la encuesta realizada en la vía pública a agentes en circulación conduce a un mayor grado de incertidumbre en los datos, debido a la demora en que incurre el agente. Existen experiencias de encuestas realizadas en hogares, a través de medios postales, de forma automática o semiautomática en la detección de etiquetas ubicadas en el vehículo, entre otras.

En el momento de la redacción del presente documento, los autores evalúan las diferentes posibilidades, volcando un especial interés en la realización de las encuestas en el lugar laboral del agente, teniendo en cuenta la supervisión y asistencia del organismo responsable de la distribución. Esta metodología se considera adecuada para organismos dependientes de la administración pública, también interesados en los resultados que pueda arrojar la encuesta, con lo cual se logra un natural apoyo en estudio.

Una vez recolectada la información, utilizando uno o varios de los métodos indicados, se procede al procesamiento de los datos (tabulación, sumarización, etc). Entre los procesos automatizados se cuenta la transformación de las direcciones origen y destino en pares de coordenadas, y la inclusión del código de zona a la cual pertenecen. A continuación surge una nueva información volcada en matrices origen-destino. Tales matrices conforman estructuras de datos más complejas si se consideran otras variables emergentes de la encuesta.

Para poder finalmente establecer caminos entre orígenes y destinos particulares, y en forma general entre centros generadores y atractores de tránsito es necesario mapear la información obtenida a través de las encuestas sobre el mapa definido por el grafo dirigido. Es en este momento cuando se consolida la información estática y dinámica.

El uso de GPS complementa esta recolección, convirtiéndose en una herramienta imprescindible para la geolocalización,

5 AUTOMATIZACIÓN DEL AFORO

El estudio realizado a través de las encuestas origen-destino se complementa con el aforo vehicular *in-situ* sobre área de interés.

Para lograr mayor velocidad en la adquisición y calidad en el almacenamiento se realiza la programación de aplicaciones que son ser utilizadas *in situ* por los recolectores de dato. Para ello, los autores han desarrollado una aplicación móvil bajo sistema operativo Android (ver figura 4).



Figura 4. Aplicación Móvil. Permite la recolección de datos in-situ, y la transmisión automática al servidor.

3. RESULTADOS OBTENIDOS

El trabajo posibilitó la adopción de una base de datos pública, y proveer mecanismos para la evaluación de la calidad de los datos provistos por la misma. A partir de esto, la transformación de dichos datos a partir del archivo XML provisto por OSM, *mapeándolos* en un formato más adecuado para el procesamiento, permitió la creación de un modelo de datos orientado a objetos, el cual es almacenado utilizando una base de datos relacional. Los datos tabulados en la base de datos permitieron al grupo un procesamiento rápido, a través del cual se pudo realizar un procedimiento de filtrado, y así apartar datos de interés: esquinas y caminos datos a través de la secuencialidad de las esquinas. Por otro lado, se ha avanzado en la programación de aplicaciones móviles para la automatización en la toma de datos.

4. FORMACION DE RECURSOS

Se trabaja en el departamento de Ingeniería en Sistemas de Información de la UTN con alumnos del tercer año en temas relacionados al modelado y simulación de flujo vehicular. Con esto se busca promover las actividades de investigación y desarrollo más allá de las actividades curriculares pertenecientes a la carrera.

5. BIBLIOGRAFIA

[1] Franco, Ramiro; Monetti, Julio, León, Oscar. Modelado y Simulación Vehicular: Proceso Gradual en la Creación de un Simulador de Tránsito. CONAIISI. 2015. Buenos Aires – Argentina. Sesión de Posters.

[2] Navarro, M; Sbriglio Franco; Monetti, Julio. ENIDI 2017. Open Street Maps: Análisis de una Base de Datos Pública y Propuesta de una Estructura de Datos para la Simulación.

[3] Slinn, M; Matthews, P & Guest, P. Traffic Engineering Design. Principles and Practice.

Segunda Edición. Elsevier. 2005. Italia. ISBN 0 7506 5865 7.

[4] Cal, R., Reyes, M.: Ingeniería de Tránsito. Fundamentos y Aplicaciones 7^a Edición. Alfaomega. (1995). ISBN: 970-15-0109-8.

[5] Schöbel, A.: Optimization in Public Transportation: Stop Location, Delay Management and Tariff Design in a Public Transportation Network. Georg-August University (Göttingen-Alemania). Springer. USA, (2006). ISBN: 978-0-387-32896-6.

Utilización de tecnología RFID/NFC para el desarrollo de un sistema de llenado selectivo de envases. Industria 4.0.

Lupi, Daniel^{1,2}; Zaradnik, Ignacio¹; Dominguez, Facundo¹; Rodriguez, Carlos¹; Kumvich, Augusto¹; Slawiski, Javier¹; García, Federico¹.

1Laboratorio de Inteligencia Ambiental Departamento de Ingeniería e Investigación Tecnológica, Universidad Nacional de la Matanza. Buenos Aires, Argentina; 2Fundación Argentina de Nanotecnología. Buenos Aires, Argentina

izaradnik@gmail.com

RESUMEN

Se presenta un sistema de llenado selectivo de envases basado en tecnología RFID/NFC (Identificación por Radio Frecuencia / Comunicación de Campo Cercano). Con el mismo se pretende ejemplificar el uso de dispositivos de tecnología de identificación por radiofrecuencia en la trazabilidad de productos manufacturados. El sistema está formado por un dispositivo electroneumático, un controlador lógico programable, PLC por sus siglas en inglés, y por una computadora. Se detalla el hardware empleado, así como el programa del PLC y el software desarrollado.

Palabras Clave: Identificación por Radio Frecuencia (RFID), Comunicación de Campo Cercano (NFC), Trazabilidad, Controlador lógico programable (PLC), Industria 4.0.

CONTEXTO

En el marco del Laboratorio de Inteligencia Ambiental del Departamento de Ingeniería e Investigación Tecnológica de la Universidad Nacional de la Matanza se viene trabajando desde hace algunos años en aplicaciones de Internet de las Cosas (IoT) [1][2][3]. El presente trabajo es parte de lo realizado en el marco del proyecto “Utilización de electrónica impresa para el desarrollo de sistemas de seguimiento, identificación y trazabilidad de productos manufacturados”, el cual se viene desarrollando en los últimos dos años [4][5][6]. Este trabajo se desarrolla con fondos pertenecientes al programa de Investigaciones PROINCE y a un subsidio PICTO del ministerio de Ciencia Tecnología e Innovación Productiva.

1. INTRODUCCION

Avances en distintas áreas tecnológicas, como los sistemas Ciber-físicos, Internet de las Cosas, Computación en la nube y el análisis de datos, entre otras, están impulsando la industria a una nueva era, la cual es denominada Industria 4.0 [7][8][9]. Esta nueva industria plantea reconvertir los procesos de fabricación para que funcione como un servicio que opera, en gran medida, como la computación en la nube. Esto implica la creación de redes y la integración de distintas compañías a la cadena de valor. Estas redes permitirían a las fábricas compartir activos, así como también especificaciones, planes e inclusive el control. Las compañías podrían contratar una fábrica por solo una cantidad de trabajo que ellas necesiten. Las fábricas distribuirían entonces el trabajo de forma de producir las cosas de mayor calidad, en la menor cantidad de tiempo y por el menor costo [10]. Esta transformación requiere una gran inversión en nuevas tecnologías [10] y aspectos asociados a ellas: seguridad [11], interfaces de usuario [12], confiabilidad [13], etc.

Jiafu Wan et al. [9] plantean que la Industria 4.0 posee dos temas principales: el primero de ellos son las fábricas inteligentes, el cual tiene como objetivo estudiar el sistema de producción inteligente y lograr instalaciones de producción distribuidas en red. El otro es la manufactura inteligente, principalmente relacionada con la gestión de la logística de producción completa de la empresa, la interacción hombre-computadora y las tecnologías de impresión 3D en procesos industriales. El trabajo que se presenta a continuación involucra aspectos de ambos temas: el llenado selectivo asociado a la

fabricación inteligente y la identificación por RFID/NFC a la manufactura inteligente.

2. LINEAS DE INVESTIGACION y DESARROLLO

El desarrollo del sistema de llenado selectivo es una parte de un sistema más complejo, el cual busca ejemplificar el uso de dispositivos RFID/NFC, realizados a través de electrónica impresa, en la trazabilidad de productos manufacturados en toda la cadena de suministros, es decir desde el ingreso de la materia prima para la fabricación hasta el usuario final.

3. RESULTADOS OBTENIDOS

3.1. Descripción General

Cada uno de los envases a ser llenados contará con un tag NFC, el cual lo identificará y permitirá grabar los datos de trazabilidad asociados al proceso de manufactura. La figura N°1 presenta una imagen del dispositivo electroneumático encargado de realizar el llenado. El mismo está formado por una cinta transportadora, un sensor óptico para la detección del envase a ser llenado, dos lectores/grabadores de NFC (uno para identificación del envase y otro para la grabación de los datos de trazabilidad del proceso) y 3 pistones neumáticos para el posicionamiento y el llenado del envase.



Figura N°1. Dispositivo electroneumático.

El control de dicho dispositivo se lleva a cabo por medio de un PLC de la firma Siemens, el modelo Logo (si bien dicho modelo está catalogado como Rele Inteligente, a los fines

prácticos de este informe se lo considera un PLC). La computadora con la cual cuenta el sistema, a través de un software realizado a medida, está encargada del manejo de los lectores/grabadores de NFC y de indicar al PLC si debe realizar el proceso de llenado o no, lo cual define en función de la lectura del tag NFC en el envase. La figura N°2 presenta una imagen del tablero en donde se encuentra montado el PLC. La figura N°3 muestra una imagen de la pantalla principal del software desarrollado, del cual se detallarán algunos aspectos en la siguiente sección.



Figura N°2. Tablero PLC.

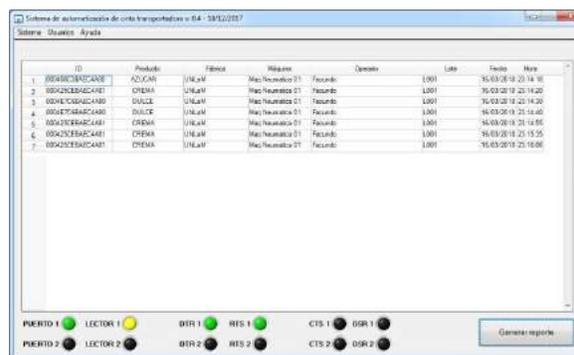


Figura N°3. Pantalla principal del software.

3.2. Desarrollo

A raíz que la implementación del sistema se realizó en simultáneo con las tareas de diseño y fabricación de los tags con electrónica impresa, es que para el mismo se emplearon tags comerciales. Dichos tags están basados en el chip NTAG203 de NXP [14], compatible con los tags Tipo 2 de NFC Forum. Los lectores empleados para la aplicación están basados en el chip PN532 de NXP [15], el cual soporta los estándares ISO/IEC 14443A. La figura N°4a

muestra una imagen de uno de los tag empleados, mientras que la figura N°4b muestra una imagen del lector. Para la comunicación con la computadora se empleó la interfaz UART del dispositivo lector.



Figura N°4a. Tag NFC. N°4b Lector NFC.

El software desarrollado posee las siguientes características:

- Acceso a través de usuario y contraseña.
- Configuración de parámetros de producción (fábrica, máquina y lote)
- Alta y baja de productos al sistema
- Configuración de puertos de comunicación con lectores
- Pantalla principal para monitorear la actividad
- Estructura de base de datos para registros de producción
- Generación de informes

Ejecutada la aplicación una pantalla para ingreso de usuario y contraseña será mostrada. Ingresados y validados dichos datos una nueva pantalla para la configuración de los parámetros de producción aparecerá. Cargados los mismos serán almacenados en un registro de la base de datos, para realizar la trazabilidad, y la pantalla principal se disparará para controlar y monitorear la producción.

En el presente desarrollo el control de la producción se refiere tan solo a la decisión de llenado del envase, en función de la identificación del mismo, y la posterior grabación de datos de trazabilidad. La decisión que tome el software generará un cambio de estado de una de las líneas de control de flujo del puerto serie. Dicha señal, previamente acondicionada, ingresará a una de las entradas del PLC para que éste realice o no el llenado. Finalizado el proceso, la computadora a través del segundo lector grabará en el tag los datos de

trazabilidad asociados, entre ellos fecha, hora, operario, máquina, etc.

Para el desarrollo de la aplicación de la computadora se empleó la versión de prueba de LabWindows/CVI [16], el cual es un entorno de desarrollo integrado ANSI C de National Instruments que incluye herramientas de ingeniería con bibliotecas integradas para análisis y diseño de UI (interfaces de usuario). Como base de datos se empleó MySQL [17], el cual es un sistema de bases de datos relacional, multihilo y multiusuario con licencia GNU GPL (si bien la aplicación presentada en este trabajo no presenta una complejidad elevada, se prevé la realización de un sistema más complejo, para el cual será necesario).

Para la realización de la programación del PLC se analizaron las distintas entradas y salidas involucradas en el proceso y el funcionamiento general del dispositivo electroneumático. Con estos datos se realizó un diagrama en bloques del funcionamiento y se llegó a la conclusión de realizar la programación de manera secuencial. La misma divide el proceso en pasos, donde cada uno de ellos depende de la finalización del paso anterior. El programa fue realizado a través de diagramas de funciones (FUP de la palabra alemana *Funktionsplan*), programación basada en bloques lógicos del tipo AND y OR, ya que, al programador le ofrece rapidez y agilidad en dicho proceso. Para su posterior documentación se migró el proyecto a programación por diagramas de contactos (KOP de la palabra alemana *Kontaktplan*), ya que ofrece un orden visual más amigable para tal fin. La figura N°5 y 6 presentan imágenes de tipos de programación.

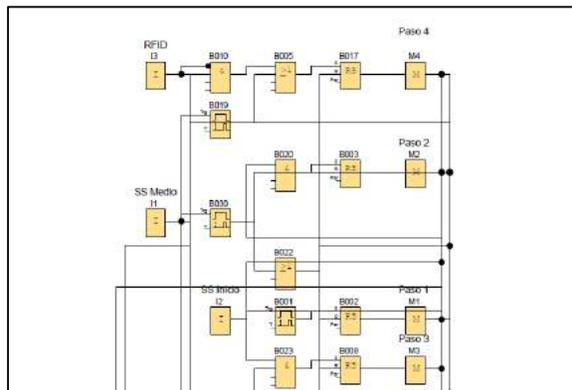


Figura N°5. Programación FUP.

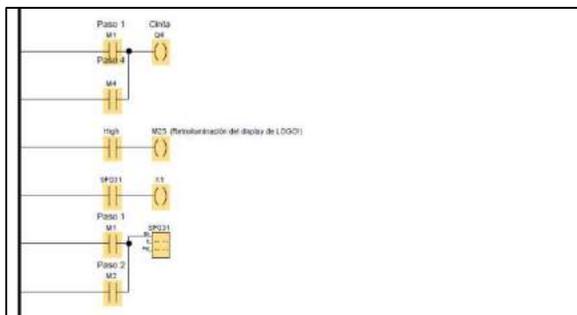


Figura N°6. Programación KOP

3.3. Conclusiones

El sistema desarrollado representa un claro ejemplo de cómo las tecnologías de RFID/NFC se pueden emplear en los procesos productivos, tanto para sumar inteligencia al proceso como para mejorar la trazabilidad del producto.

Como trabajo futuro está previsto la ampliación del sistema, el cual incluirá una etapa de etiquetado previa y una etapa de embalaje posterior. El software se ampliará a fin de incorporar estas nuevas etapas y se integrarán otros desarrollos realizados por el grupo de investigación para generar un sistema que permita la trazabilidad de toda la cadena de suministros.

4. FORMACION DE RECURSOS HUMANOS

El ámbito de este proyecto permitió tanto la formación grupal del equipo de trabajo así como la individual de cada uno sus miembros y colaboradores. La formación grupal busco generar conocimiento en el área de la electrónica impresa y como su aplicación puede mejorar los procesos de manufactura. Los casos de la formación individual se enumeran a continuación: Ignacio Zaradnik la gestión de grupos de trabajos, Javier Slawiski la interpretación de normas y estándares, Monica Canziani la revisión bibliográfica y la elaboración de estados del arte, Facundo Domínguez el diseño de aplicaciones de software, Diego Turconi el diseño de aplicaciones de sistemas embebidos y tanto Augusto Kumvich, Federico García como Carlos Rodríguez la automatización de procesos con RFID y PLC.

5. REFERENCIAS

- [1] Brengi, Diego; Canziani, Monica; Gomez, Rodrigo; Gwirc, Sergio; Lupi Oreste; Moltoni, Andrés; Nassipián, Veronica; Slawiski Javier; Zaradnik, Ignacio, “Sistema inalámbrico de microestaciones meteorológicas para aplicaciones agropecuarias”. Congreso Argentino de Sistemas Embebidos 2013. ISBN 978-987-9374-88-77.
- [2] Canziani, Mónica; Gomez, Rodrigo; Lupi, Daniel; Nassipián, Verónica; Slawiski Javier; Turconi, Diego; Zaradnik, Ignacio, “Plataforma de conexión de Redes Eléctricas Inteligentes a Internet de las Cosas” en el Congreso Argentino de Sistemas Embebidos 2014. ISBN 978-987-45523-2-7.
- [3] Bernis, Ariel; Turconi, Diego; Benacerraf, Mario; Dominguez, Facundo; Lupi, O. Daniel; Zaradnik, Ignacio; Rzepa, Antonio, “Sistema de seguimiento de dosimetría personal”. VII congreso de microelectrónica aplicada 2016. ISBN: 978-987-733-068-7.
- [4] Canziani, Monica; Lupi, Daniel; Ortiz, Juan José; Slawiski, Javier; Zaradnik, Ignacio, “Tecnologías inalámbricas para sistemas de seguimiento, identificación y trazabilidad de productos”. VII Congreso de Microelectrónica Aplicada 2016 (uEA 2016), San Luis, Argentina. ISBN 978-987-733-068-7.
- [5] Lupi, Daniel; Zaradnik, Ignacio; Turconi, Diego; Dominguez, Facundo, “Sistema de visualización de precios para supermercados”. Congreso Argentino de Sistemas Embebidos 2017 (Case 2017), Buenos Aires, Argentina. ISBN: 978-987-46297-3-9.
- [6] Canziani, Monica; Zaradnik, Ignacio; Tantignone, Hugo; Lupi, Daniel; Villares Had, Bernardo, “Procesos Tecnológicos, Tintas y Sustratos empleados en Electrónica Impresa”. VIII Congreso de Microelectrónica Aplicada 2017 (uEA 2017), Córdoba, Argentina.
- [7] Mario Hermann; Tobias Pentek; Boris Otto, “Design Principles for Industrie 4.0

- Scenarios”. 2016 49th Hawaii International Conference on System Sciences.
- [8] Teth Azrael Cortés Aguilar, “Los sistemas embebidos en la industria 4.0”. Sistemas embebidos: Estado actual con visión al futuro. ISBN 978-607-9394-08-0, Asociación Mexicana de Mecatrónica A.C.
- [9] Jiafu Wan; Hu Cai; Keliang Zhou, “Industrie 4.0: Enabling Technologies”. 2015 International Conference on Intelligent Computing and Internet of Things (ICIT)
- [10] David Alan Grier, “The Radical Technology of Industrie 4.0”. COMPUTER PUBLISHED BY THE IEEE COMPUTER SOCIETY. 2017.
- [11] Michael Waidner; Michael Kasper, “Security In Industrie 4.0 - Challenges and Solutions for the Fourth Industrial Revolution”. 2016 Design, Automation & Test in Europe Conference & Exhibition.
- [12] Thies Pfeiffer; Jens Hellmers, Eva-Maria Schön, and; Jörg Thomaschewski, “Empowering User Interfaces for Industrie 4.0”. Proceedings of the IEEE | Vol. 104, No. 5, May 2016.
- [13] PAULA FRAGA-LAMAS; TIAGO M. FERNÁNDEZ-CARAMÉS1; ÓSCAR BLANCO-NOVOA; MIGUEL VILAR-MONTESINOS, “A Review on Industrial Augmented Reality Systems for the Industry 4.0 Shipyard”. ACCESS.2018.2808326, IEEE Access. 2018.
- [14] NXP, “NTAG203 NFC Forum Type 2 Tag compliant IC with 144 bytes user”, Rev. 3.0 — 17 October 2011.
- [15] NXP, “PN532/C1 Near Field Communication (NFC) controller”, http://www.nxp.com/documents/short_data_sheet/PN532_C1_SDS.pdf, última visita 17/03/2018.
- [16] National Instruments, <http://www.ni.com/lwcv/whatis/esa/>, última visita 17/03/2018.
- [17] MySQL, <https://www.mysql.com/>, última visita 17/03/2018.

Realidad Virtual y Aumentada, Big Data y Dispositivos Móviles: Aplicaciones en Turismo

Feierherd Guillermo¹, Depetris Beatriz¹, Huertas Francisco¹, González Federico¹, Romano Lucas¹, Viera Leonel¹, Horas Fabiola¹, Delía Lisandro²

¹ Instituto de Desarrollo Económico e Innovación
Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur
Hipólito Irigoyen 880 - Ushuaia - Tierra del Fuego

² III - LIDI - Facultad de Informática - Universidad Nacional de La Plata

{gfeierherd, bdepetris, fhuertas, fgonzalez, lromano, lviera}@untdf.edu.ar
fabolahoras@gmail.com
ldelia@lidi.info.unlp.edu.ar

RESUMEN

Las Tecnologías de la Información y la Comunicación (TICs) son “tecnologías disruptivas”, pues permiten resolver antiguos problemas mediante soluciones impensables previamente.

La mayoría de las actividades humanas las han incorporado, produciendo la aparición de soluciones novedosas y un importante aumento de la productividad.

La actividad turística no es ajena a estas transformaciones. Alcanza con ver cómo se realizan hoy las reservas y contrataciones de los servicios turísticos para advertir los cambios introducidos para productores y consumidores.

Por otra parte, el auge de teléfonos inteligentes (smartphones), contribuyó a la ubicuidad de la computación y, con ello, a la generación de grandes volúmenes de datos (big data) y nuevos paradigmas sobre las formas en las que se utilizan estos recursos. A su vez, dispositivos cada vez más potentes y económicos facilitan tecnologías como las de realidad virtual y realidad aumentada al público masivo.

El proyecto busca relevar los usos que la industria turística está haciendo de estas tecnologías en forma individual o

combinada, para proponer alternativas de aplicación en el ámbito de nuestra provincia. A fin de demostrar la factibilidad de las propuestas se propone desarrollar algunas aplicaciones experimentales.

Palabras clave: Realidad virtual; Realidad Aumentada; Big Data; Dispositivos Móviles; Turismo; Smart Destinations

CONTEXTO

El presente proyecto forma parte del “Área de Investigación 8 - Desarrollo Informático”, del Instituto de Desarrollo Económico e Innovación de la Universidad Nacional de Tierra del Fuego (UNTDF).

El proyecto se presentó a la convocatoria realizada por la UNTDF en septiembre de 2016 y resultó aprobado luego de haber sido sometido a evaluación externa. (Resolución Rectoral 060/2017 del 10/04/2017).

El financiamiento requerido es suministrado por la UNTDF.

El proyecto se desarrolla formalmente desde el 01/03/2017 hasta el 28/02/2019.

INTRODUCCIÓN

El rápido desarrollo de las Tecnologías de la Información y la Comunicación (TICs) produce cambios en todos los niveles organizacionales, tanto en el ámbito público como privado.

Entre las transformaciones que interesan al proyecto están las vinculadas al concepto de Destinos Turísticos Inteligentes (STD por Smart Tourism Destinations), el que constituye una derivación directa del concepto de Ciudades Inteligentes (Smart Cities).

Si bien es difícil encontrar definiciones que estén universalmente aceptadas, es posible decir que una ciudad inteligente es aquella en la que las TICs se introducen estratégicamente buscando mejorar la competitividad de la ciudad y, al mismo tiempo, la calidad de vida de sus ciudadanos. [1]

Por su parte, un Destino Turístico Inteligente puede ser definido de distintas maneras. Como lo señala Alfonso Vargas-Sánchez [2] en su revisión de la literatura sobre el tema, un STD es definido por López de Ávila como *“un destino turístico innovador, construido sobre la infraestructura de tecnología actualizada, garantizando el desarrollo sostenible de las áreas turísticas, accesible para todos, facilitando la interacción de los visitantes y su integración con el entorno, incrementando la calidad de la experiencia en el destino y mejorando la calidad de vida de los residentes”*. A su vez señala Vargas-Sánchez que Gretzel, Sigala, Xiang & Koo consideran que el turismo inteligente es el *“turismo apoyado por esfuerzos integrados en un destino para recopilar y aprovechar los datos derivados de la infraestructura física, las conexiones sociales, las fuentes gubernamentales y organizativas en combinación con el uso de tecnologías*

avanzadas para transformar esos datos en experiencias in situ y propuestas de valor comercial con un enfoque claro en la eficiencia, la sostenibilidad y el enriquecimiento de la experiencia.”

Como surge de las definiciones anteriores es evidente que las TICs son facilitadoras. Como bien señalan Boes, Buhalis e Inversini, *“los destinos pueden desarrollar su inteligencia alineando las dimensiones clave de liderazgo, capital social, innovación y capital humano, utilizando las TICs como la “info-estructura” que facilita la co-creación de valor / experiencias para sus visitantes y competitividad para su industria [3]. Al mejorar la inteligencia de las dimensiones de Ciudad Inteligente (personas, vida, movilidad, medio ambiente, economía y gobierno), los destinos crean las condiciones para apoyar el desarrollo de Destinos de Turismo Inteligente donde se prioriza la interconexión, co-creación y la creación de valor, a través de la implementación de aplicaciones tecnológicas e infraestructuras TIC como Cloud Computing e Internet de las Cosas [4]. Las sinergias entre el interés y la preferencia garantizan que todas las partes interesadas se beneficien del proceso y que se desarrollen mejores experiencias y calidad de vida para todas las partes involucradas en el destino turístico.*

La tecnología está inmersa en prácticamente cualquier organización o entidad, los destinos turísticos van a potenciar las sinergias entre la tecnología ubicua y distintos componentes sociales para proveer experiencias enriquecedoras a los turistas, ya sea antes, durante o luego de su viaje. Como ha ocurrido en todos los ámbitos en los que criteriosamente se introduce tecnología, los destinos que hagan un buen uso de éstas podrán incrementar sus niveles de competitividad.

Las TICs hacen que las ciudades sean más accesibles y disfrutables, tanto para residentes como para turistas, gracias a servicios interactivos que interconectan distintos niveles de gobierno con empresas y proveen información en tiempo real de utilidad para todos ellos. Además, los datos resultantes pueden ser analizados por los organismos involucrados para el desarrollo de mejores políticas.

Desde una perspectiva turística, las TICs pueden contribuir generando valor agregado a las experiencias de los turistas y a la vez mejorar la eficiencia de las organizaciones relacionadas, facilitando la automatización de algunos procesos o la obtención de información valiosa. Es un hecho que las tecnologías de la información y la comunicación (TICs) han significado una innovación disruptiva en las relaciones entre oferta y demanda a nivel de turismo.

El entorno digital permite promocionar destinos, productos y servicios y, a la vez, conocer a los turistas como nunca antes se había hecho, ofreciéndoles experiencias personalizadas y de mayor calidad: segmentación e hipersegmentación del mercado, reducción de costos, mayor eficiencia y competitividad, etc.

Desde este punto de vista, las TICs han abierto un nuevo horizonte en el sector turístico, que plantea retos y oportunidades, y que requiere un gran esfuerzo de adaptación, tanto por parte de las empresas que prestan estos servicios, como por la gestión del destino liderada por el sector público.

La democratización de la tecnología ha sido la causa de que el sector haya pasado de estar controlado por la oferta (de las empresas) a estar dominado por la demanda (de la sociedad civil).

Por otra parte ha cambiado la forma de acceder a internet: se ha pasado de la

computadora de escritorio y la notebook a los dispositivos móviles (tablets y smartphones principalmente). A su vez, estos nuevos dispositivos ya no solo se usan para comunicarse entre personas, sino que poseen una serie de sensores y accesorios que los convierten en aparatos “todo en uno” (cámara de fotos, filmadora, gps, medio de pago, etc.) En consecuencia, tanto el sector público como el privado tienen que estar preparados para cubrir los requerimientos y necesidades del turista en el siglo XXI: más informado (a través de webs, redes sociales, etc.), que organiza su viaje personalmente y que requiere una serie de servicios digitales en el destino, similares a los que tiene en su residencia habitual (por ejemplo, conectividad inalámbrica para todos sus dispositivos).

La llamada “Internet de las cosas” (IoT por Internet of Things) está ganando terreno rápidamente entre las TICs. El término fue propuesto en 1999 por Kevin Ashton -investigador del MIT- quien definió IoT como una red que conecta cualquier cosa en cualquier momento y lugar, para identificar, localizar, administrar y monitorear objetos inteligentes.

La idea detrás de IoT es generar interacciones automáticas en tiempo real entre distintos objetos conectados a internet. Así, IoT facilita que distintas plataformas puedan transmitir datos obtenidos mediante diferentes sensores y que esos datos puedan ser procesados e interpretados en tiempo real.

En un contexto turístico, los turistas pueden usar sus teléfonos móviles para explorar el destino y sus eventos de interés, utilizando información provista por los gobiernos, agencias privadas e incluso otros turistas o ciudadanos.

Todas estas actividades llevadas a una escala masiva producen una cantidad de información multidimensional

conocida como Big Data.

Haciendo uso del Big Data, las organizaciones turísticas pueden extraer información valiosa para mejorar la experiencia de los usuarios, proponiendo una nueva forma de relación con los turistas. Los destinos turísticos que primero aprovechen estas condiciones, harán una importante diferencia respecto de los que queden rezagados.

Por su parte, la Realidad Virtual (VR) permite a los turistas tener una experiencia inmersiva de los destinos, aún a distancia, mostrándole a la persona “cómo es y cómo se siente el lugar propuesto”. La VR facilita a los turistas visualizar y recorrer distintos atractivos, obteniendo información complementaria en formatos multimediales (texto, audio, imagen, video) en un entorno 3D o de 360°.

A su vez, la Realidad Aumentada (AR) puede utilizarse para mostrar cómo se veía un mismo atractivo en distintas épocas (por ejemplo, cómo se veía Ushuaia hace 100 años) o agregar, en tiempo real, información de interés a un elemento en cuestión.

Con Realidad Virtual y Realidad Aumentada la promoción de los destinos turísticos se vuelve una atracción en sí misma, incrementando el interés de las personas y aumentando así el mercado potencial de turistas.

Es deseable entonces contar con espacios turísticos consolidados sobre la base de una infraestructura tecnológica de vanguardia, incluyendo sistemas inteligentes que obtengan la información de forma automática, la analicen y comprendan los acontecimientos en tiempo real, facilitando la toma de decisiones y la interacción del visitante con el entorno turístico.

La innovación, el aprovechamiento de las nuevas tecnologías, la sostenibilidad ambiental y económica,

como así también la cooperación público-público y público-privada, son condiciones sine qua non para el desarrollo turístico actual, donde el gobierno, los ciudadanos y las empresas locales deben asociarse con los turistas para comprenderlos, ofrecerles lo mejor y aprender de ellos.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Como ha sido mencionado previamente los temas de investigación / desarrollo / innovación para el proyecto son los siguientes:

- Realidad Virtual
- Realidad Aumentada
- IoT
- Big Data

especialmente en su vinculación con dispositivos móviles.

RESULTADOS Y OBJETIVOS

El objetivo del proyecto es analizar el estado del arte en las tecnologías de (Realidad Virtual [VR], Realidad Aumentada [AR] y Big Data [BD]), poniendo especial atención al uso que se hace de las mismas en el dominio del turismo, con el fin último de proponer y desarrollar prototipos de aplicaciones enfocadas al turismo local.

En ese marco se realizaron (durante este primer año) las siguientes actividades:

- 1.- Encuesta global que explora las condiciones que deberían caracterizar a un sistema digital de información turística en áreas protegidas para las generaciones de nativos digitales. Participaron 269 personas de 12 países. [5]
- 2.- Evaluación de los smartphones disponibles en el mercado argentino y sus capacidades para ofrecer contenidos en

VR y AR, con el objetivo de establecer qué funciones podrían esperarse actualmente de una app que considere a la gran mayoría de los usuarios.

3.- Identificación automática de flora mediante AR e inteligencia artificial (computer vision). Se compararon las plataformas IBM Watson y Microsoft Azure para el reconocimiento de especies nativas de árboles, con excelentes resultados en los dos casos.

4.- Se realizaron pruebas de filmaciones en 360° aplicando técnicas de post-producción para eliminar trípodes y objetos que no deberían verse en escena.

5.- Se diseñó e imprimió en 3D un soporte para colgar una cámara 360° de un dron. Se probaron distintos diseños hasta obtener uno que tuviera la menor vibración posible y se ajustó la estabilización con el giroscopio de un teléfono iPhone conectado a la cámara.

6.- Se filmó en Antártida y se produjeron videos en 360° que serán publicados en un nuevo sitio web con información turística oficial de la provincia.

7.- Parte de los videos mencionados se utilizaron para la producción de los contenidos en VR para el centro de interpretación Antártica que abrirá sus puertas en Ushuaia a partir de abril de 2018. [5]

Durante los próximos meses se espera realizar pequeñas apps experimentales, principalmente para dar seguimiento al reconocimiento de especies y para la producción y recolección de información en formato de Big Data.

Respecto al primer tema se espera desarrollar una app de AR para reconocer en tiempo real las especies de árboles de Tierra del Fuego, principalmente las que se encuentran en la zona sur de la provincia.

En relación a Big Data se desarrollará una app para capturar de manera anónima

la información de GPS de los turistas, con lo que luego se espera: estimar turistas en un lugar en tiempo real; dibujar mapas con los recorridos preferidos por la gente; conocer desde qué ciudades vienen y hacia cuáles van los turistas; listar los alojamientos y restaurantes más populares; hacer un ranking de puntos de interés en el destino

FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está compuesto por tres Profesores Titulares (dos vinculados a la Licenciatura en Sistemas y uno a la Licenciatura en Turismo), dos Asistentes, dos Alumnos (uno de ellos también Asistente Alumno) y un Profesor Adjunto externo (UNLP), también vinculado a las TICs.

Recientemente, Federico González, finalizó su Máster en Ciudades Inteligentes en la Universidad de Girona, España. Su tesis “Herramientas digitales de información turística en Áreas Protegidas para Millennials y Gen Z. Un caso de realidad virtual y aumentada en viajes antárticos”, está íntimamente ligada al proyecto. [6]

Lisandro Delia finalizó la carrera de Especialista en Ingeniería de Software en la Universidad Nacional de La Plata. Su trabajo titulado "Desarrollo de Aplicaciones Móviles Multiplataforma" [7] también tiene aplicación directa al tema del presente proyecto.

Leonel Viera finalizó la Licenciatura en Informática en la Universidad Nacional de la Patagonia San Juan Bosco. Su tesina consistió en diseñar un “Marco de trabajo para el desarrollo de aplicaciones móviles híbridas”. [8]

Cabe mencionar también que Francisco Huertas, integrante del proyecto vinculado a la Licenciatura en Turismo dirige el “Relevamiento y puesta

en valor del Patrimonio Artístico (bustos, composiciones, monolitos, placas, estatuas y mástiles) en espacios públicos de la ciudad de Ushuaia”, proyecto presentado y aprobado en la misma convocatoria que el que aquí se describe.

REFERENCIAS

- [1] Andrea Caragliu, Chiara Del Bo, and Peter Nijkamp. (2017, 03). Smart Cities in Europe. Journal of Urban Technology. [Online]. 18(2). Disponible en https://www.researchgate.net/publication/46433693_Smart_Cities_in_Europe
- [2] Alfonso Vargas-Sánchez. (2017, 03). Exploring the concept of Smart Tourist Destination. Enlightening Tourism. A Pathmaking Journal. [Online] 6(2). Disponible en https://www.academia.edu/30399967/Vol_6_No_2_2016_July-December
- [3] Kim Boes, Dimitrios Buhalis and Alessandro Inversini. (2017, 03). Conceptualising Smart Tourist Destination Dimensions. Disponible en https://www.researchgate.net/profile/Dimitrios_Buhalis/publication/272576525_Conceptualising_Smart_Tourism_Destination_Dimensions/links/54e9d4390cf25ba91c7ff25c.pdf
- [4] Taewo Nam y Theresa Pardo, (2017, 03), “Conceptualising Smart City with Dimensions of Technology, People and Institutions”, presentado en 12th Annual International Conference on Digital Government Research, College Park, MD, USA, Junio 12 - 15, 2011. Disponible en https://inta-ai.vn.org/images/cc/Urbanism/background%20documents/dgo_2011_smartcity.pdf
- [5] Presentación de la Experiencia Antártica, Tierra del Fuego, Argentina. <https://www.youtube.com/watch?v=enMDsmxYktk>
- [6] Tesis Final de Máster en Smart City, Federico González, Universidad de

Girona. <https://goo.gl/9NXQ9i>

- [7] Trabajo Final de Especialista en Ingeniería de Software, Lisandro Delia, Universidad Nacional de La Plata. <http://sedici.unlp.edu.ar/handle/10915/60497>
- [8] Tesina de Grado de Licenciado en Informática, Leonel Viera, Universidad Nacional de la Patagonia San Juan Bosco, Sede Ushuaia. <https://docs.google.com/document/d/172OoluXn891cve80tpVVgyqi-POSJGdrM33oFEYL9Nk/edit?usp=sharing>

Dispositivo de asistencia de personas mediante monitoreo IoT

Waldo Valiente, Esteban Carnuccio, Mariano Volker, Graciela De Luca, Gerardo García, Daniel Giulianelli, Raúl Villca, Marcos Vittorio
Departamento de Ingeniería e Investigaciones Tecnológicas

Universidad Nacional de La Matanza

Dirección: Florencio Varela 1703 – CP 1754 – {wvaliente, ecarnuccio, mvolker, gdeluca, ggarcia, dgiulian}@unlam.edu.ar, raul.villcasd@gmail.com, marcosvittorio@hotmail.com

RESUMEN

Los objetos y las personas, gracias a la tecnología actual, pueden ser comunicados y utilizados para mejorar las condiciones de vida. Este trabajo intenta hacer llegar la tecnología para el cuidado de las personas, realizando un aporte social, ya que se trata de desarrollar un dispositivo simple, amigable y de bajo costo. Será simple debido a que solo realizará funciones para el monitoreo y comunicación del estado de la persona. Amigable, ya que será fácil de usar para el monitoreo y contará con una aplicación Android en un dispositivo del cuidador. Los datos obtenidos durante el monitoreo serán almacenados en un servidor en la nube pudiendo ser procesados y utilizados posteriormente de ser necesario. Contará con el sistema de caídas desarrollado por este equipo de investigación, adaptándolo a un dispositivo pequeño y con baterías de larga duración. De bajo costo, debido a una producción en serie mediante circuitos y software a medida.

Palabras clave: *Computación en la nube, securización de datos, monitoreo de personas, geolocalización, IoT, Android, giroscopio, acelerómetro.*

CONTEXTO

Nuestra Línea de Investigación es parte del proyecto *Dispositivo de asistencia de personas mediante monitoreo y análisis de datos en la nube*, dependiente de la Unidad Académica del *Departamento de Ingeniería e Investigaciones Tecnológicas*, perteneciente al programa de Investigaciones PROINCE de la Universidad Nacional de La Matanza, el cual es formado por docentes, investigadores y alumnos de las carreras de ingeniería en informática e ingeniería en electrónica. Este proyecto es continuación de los trabajos que viene realizando el grupo de investigación, en sistemas operativos, computación de alto rendimiento, entre otros, en el área de Internet de las cosas (IoT).

1. INTRODUCCIÓN

Los grandes avances que ha sufrido la

tecnología en los últimos años han influido considerablemente en la vida cotidiana de las personas de cualquier edad. Hoy en día, niños, adolescentes y adultos emplean distintos elementos tecnológicos para estar interconectados entre sí, compartiendo información involuntariamente. Aprovechando esta característica evolucionó el concepto de Internet de las Cosas, generando relaciones entre los objetos utilizados en la vida diaria que antes era impensados realizar, produciendo una gran cantidad de datos que se puede analizar a través de Big Data [1]. Por consiguiente, existen diferentes disciplinas en las que se puede implementar sistemas que hagan uso de IoT. En este sentido es de suma importancia la aplicación de la tecnología para el cuidado de niños, personas mayores, enfermos, o con capacidades diferentes. Siendo de vital importancia la asistencia rápida y oportuna, ante determinadas situaciones que puedan sufrir esas personas. Tratando de esta forma de prevenir problemas mayores de salud. En ese contexto, en la investigación anterior [2], se implementó el prototipo de un sistema para detectar las caídas de personas mayores de edad, utilizando las placas Intel® Galileo Generación I y el sensor MPU6050. El sistema de detección de caídas está compuesto por distintos programas, que se muestran en la Fig. 1, donde cada uno de ellos realiza una funcionalidad específica.

El detector de caídas en el sistema embebido [2], ejecuta un programa que detecta caídas de la persona mayor que lo esté utilizando. Su programación emplea un algoritmo que recolecta y analiza la información motriz diaria de la persona adulta, con el fin de detectar la ocurrencia de cualquiera de los siguientes eventos. Estos sucesos son una mejora de los propuestos por [3] y [4] en sus trabajos.

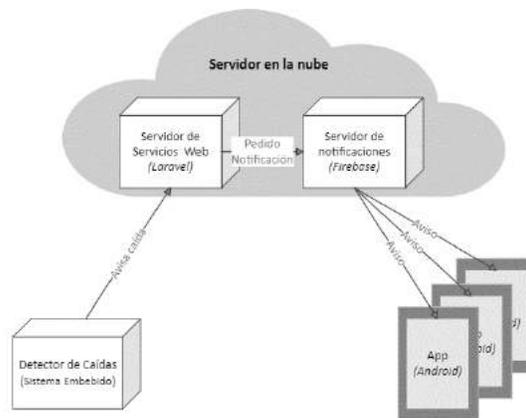


Fig. 1-Componentes del sistema detector de caídas.

El primero de ellos ocurre en el momento que se haya producido una caída, el segundo al transcurrir un determinado tiempo de haberse caído la persona y no haya conseguido ponerse de pie. Finalmente, un tercero cuando se reincorpore. Cuando el programa detecte algunos de estos eventos, automáticamente enviará mensajes de alertas a una aplicación Android, que estará instalada en el teléfono inteligente de la persona a cargo del mayor. Para realizar este cometido, se utiliza un servidor en la nube como intermediario para la transferencia de datos.

Con la experiencia lograda en la investigación anterior [2], se advierte la necesidad de evolucionar el dispositivo detector de caídas. Como anexo se agregarán capacidades que permitirán obtener parámetros biométricos simples y de ubicación, para una rápida asistencia. Con la evolución de los teléfonos inteligentes, sus sensores y la facilidad para conectarse, desarrollar una solución con este tipo de tecnología parece tentadora. Sin embargo en la investigación se considera que esta opción no resulta viable por tres factores. El primero por el marco de nuestra economía, los equipos móviles tienen un alto costo. El segundo algunas personas no suelen estar familiarizadas con esta tecnología y lo ven

como algo extraño e intrusivo. El tercer factor requiere constante intervención humana, no siendo un dispositivo dedicado para tal fin. Por eso se pretende desarrollar un dispositivo a medida con tecnologías asentadas que son utilizadas en nuestro país. Esto permite que sea de bajo costo y simple de utilizar tanto [5].

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

El apartado anterior da cuenta de la obtención de un producto estable que cumple con las funcionalidades propuestas, incluso superando a proyectos en una línea de investigación similar. Sin embargo al plantearse la viabilidad de convertirlo en un producto de uso diario se detectaron ciertas falencias. En este sentido las limitaciones que se destacan en el prototipo desarrollado, son el consumo de batería, el tamaño, su necesidad de depender de conexión de Wifi y la falta de datos biométricos. Por otro lado resulta necesario que el servidor permita emplear mecanismos de securización en los datos intercambiados hacia la nube, para mayor seguridad de la información privada de los usuarios. Además es necesario que la aplicación del teléfono móvil muestre mayor información detallada y crezca en funcionalidades en concordancia a lo anteriormente mencionado.

Mejoras al prototipo

Con respecto al consumo de batería resultante, según la especificación del producto [6], las placas Intel® Galileo GEN 1 requieren un voltaje de 5 volts y tiene un consumo de 800mA (Estos consumos no incluyen los periféricos como sensores y Wifi). No obstante las baterías recargables y portables que pueden producir esos niveles de tensión requieren de un tamaño superior a la de las pilas AA. La carga útil de estas baterías apenas logran cubrir el uso de pruebas con el dispositivo, ya que requieren ser

recargadas constantemente, cada dos horas aproximadamente. Por consiguiente lo que se busca reducir es el consumo del dispositivo, cambiando la arquitectura. Por ese motivo se están investigando varios modelos, por ejemplo los dispositivos de la familia Cortex-M3®, que requieren de un voltaje que puede llegar a los 3.6 v y el consumo puede rondar entre los 80 a 150 mA [7] [8]. Otras tecnologías como ARM nano® tiene voltajes de referencia de 3.3v y un consumo de 50mA [9].

La primera versión del prototipo presentaba elevadas dimensiones y no resultaba cómodo para ser utilizado de modo diario, debido a que es necesario llevarlo en la cintura (Fig. 2 - A). En consecuencia, al migrar de arquitectura se busca conseguir un menor tamaño del producto final. La ubicación del dispositivo está condicionada al diseño del algoritmo de caídas, donde se espera que éste se encuentre en una posición fija y vertical en el cuerpo del usuario, ya sea en su cintura o pecho, descartando así lugares como muñeca o tobillos. Esta ubicación, en el algoritmo, da un eje fijo de referencia ya que siempre se encuentra afectado por la gravedad. Por consiguiente el nuevo prototipo debería estar en la cintura. A modo de ejemplo la (Fig. 2 – B) muestra las nuevas proporciones esperadas para la siguiente generación del producto. Reduciendo así el tamaño a menos del 20% del prototipo actual.

Cabe aclarar que más adelante en la etapa final del proyecto, se va a desarrollar un circuito impreso, diseñado a medida, que miniaturice aún más el producto final.

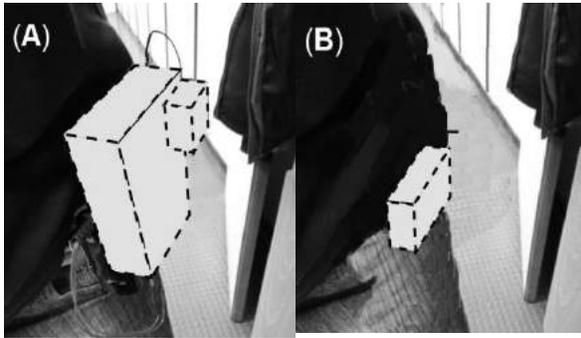


Fig. 2 Dimensiones del prototipo actual (A). Dimensiones esperadas del nuevo prototipo (B).

Con respecto a la forma de conectarse a la nube, se planteó la necesidad de que el dispositivo esté conectado más allá de Wifi, que es de uso interno. Por esto se utilizará conexión por red GSM, utilizando solo los paquetes de datos 3G/4G.

Como datos biométricos se están evaluando la implementación de sensores que permitan el monitoreo del estado de una persona. Como por ejemplo pulsometro, de temperatura, de ritmo cardiaco, entre otros. Junto con datos de la localización de la caída obtenida por medio de un GPS.

Mejoras al Servidor web

El dispositivo embebido al recolectar mayor cantidad de información y al ser datos sensibles de la persona monitoreada, surgen dos requisitos a agregar al servidor web. El primero con respecto al almacenamiento de la información, solo con fines de una posterior revisión. Mientras que el segundo, es en cuanto a la encriptación o cifrado en la transferencia de los datos.

Mejoras a la aplicación móvil

La aplicación móvil actualmente permite realizar la gestión de personas a monitorear y la recepción de los sucesos reportados por el dispositivo que realiza el monitoreo. La nueva

solución, al recolectar datos biométricos, será necesario adaptar la aplicación móvil para que realice distintas consultas y generar gráficos estadísticos. Además será necesario que la información consultada sea transferida con algún tipo de securización confiable para mayor seguridad.

3. RESULTADOS OBTENIDOS/ESPERADOS

Las mejoras que se proponen realizar en el presente trabajo de investigación, tiene como objetivo desarrollar un nuevo prototipo mejorado IoT. Dando la base para la elaboración de un producto final, que incluya las mejoras mencionadas en los párrafos previos. De esta forma se pretende diseñar un circuito impreso con todos los componentes, un servidor web securizado y aplicaciones Android masivas.

4. FORMACIÓN DE RECURSOS HUMANOS

La presente línea de investigación dentro del departamento de Ingeniería e Investigaciones Tecnológicas forma parte del trabajo que uno de los investigadores se encuentra realizando para su maestría.

Completan el grupo de investigación dos investigadores en formación y dos alumnos que se encuentran finalizando sus carreras de grado.

5. BIBLIOGRAFÍA

- [1] Bonilla-Fabela, Morales-Escobar y Guajardo-Muñoz, «IOT, El Internet de las Cosas y la Innovación de sus Aplciaciones,» 2 ed., 2016, pp. 2313-2340.
- [2] Esteban Carnuccio, Waldo Valiente, Mariano Volker, Graciela De Luca, Gerardo Garcia, Daniel Giulianelli, Sebastián Barillaro, «Desarrollo de un Prototipo

- detector de caídas utilizando la placa Intel Galileo Generación I y el sensor MPU6050,» de *IX Workshop Innovación en Sistemas de Software (WISS)*., La Plata, 2017, pp. 954-964.
- [3] R. Blanco, «Sistema de detección de caída en personas de la tercera edad para uso en centro geriátricos,» Bogota, 2010.
- [4] E. Oporto Díaz, «Diseño de un sistema inalámbrico de detección de caídas aplicado a personas de la tercera edad basado en acelerómetro y teléfono móvil,» Lima-Peru, 214.
- [5] J. Aceros, M. Cavalcante y M. Doménech, «Identidad de lugar en usuarios de teleasistencia un análisis conversacional,» Bellaterra, España, 2013.
- [6] Intel(R), «Galileo Datasheet,» de *Intel Galileo*, 2013, p. 13.
- [7] STMicroelectronics, «Datasheet - production data Ultra-low-power 32-bit MCU ARM®-based Cortex®-M3,» de *STM32L151x6/8/B STM32L152x6/8/B*, STMicroelectronics, 2016.
- [8] STMicroelectronics, «Datasheet - production data- Medium-density performance line ARM®,» de *STM32F103x8*, STMicroelectronics, 2015.
- [9] «ARM-N8LD,» de *Narrow band ISM Transceiver 868MHz / 27 Narrow band ISM Transceiver 868MHz / 27dBm*, ATIM, 2015.

Gestión del control de acceso con tecnología open source en proyectos de domótica.

<p>Mariano Emanuel Alejandro López Departamento de Informática, Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, Corrientes, Argentina. Asignatura Redes de Datos, carrera LSI. m_villa@hotmail.com</p>	<p>Leopoldo José Ríos Departamento de Informática, Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, Corrientes, Argentina. Asignatura Redes de Datos, carrera LSI. ljr@comunidad.unne.edu.ar</p>
--	---

Resumen

Esta línea de trabajo aborda el estudio y despliegue de un sistema domótico gestionado por software, a ser incorporado al Laboratorio LRDTBD del Departamento de Informática de la Facultad como parte de sus estrategias de investigación. Se pretende desarrollar sistemas de control de acceso para instituciones públicas y privadas de la región Nordeste, por ser uno de los problemas más complejos en las administraciones públicas, por diferentes motivos.

En sistemas como el que se propone, es imprescindible el control de acceso y la transferencia de servicios a consumir, para ello es necesario el despliegue de infraestructuras como Active Directory (administración segura de usuarios centralizada), API REST (transferencia de servicios web) y “JWT” (autenticación vía JSON Web Tokens), que resultan atractivas para resolver los problemas mencionados.

El proceso de autenticación utilizado en aplicaciones web, además de ser un aspecto sencillo de desplegar, ha seguido por buen tiempo un patrón común: usuario-contraseña-cookie, sin embargo, esto ha cambiado por motivos relacionados con la forma en que se construyen y distribuyen las aplicaciones modernas. Se busca que el inicio de sesión de usuarios sea realizado por una infraestructura de identidad sólida, de inicio de sesión único, con soporte para redes sociales y soporte corporativo (LDAP).

Palabras clave: Domótica, RBAC, JWT, LDAP, API REST.

Contexto

La línea de Investigación y Desarrollo presentada en este trabajo corresponde al proyecto PI-F17-2017 “Análisis e implementación de tecnologías emergentes en sistemas computacionales de aplicación regional.”, denominado Grupo de Investigación en Innovación en Software y Sistemas Computacionales (GISSC), acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2018-2021, en vinculación al Laboratorio LRDTBD del Departamento de Informática, de la FaCENA. Se pretende dotar al Laboratorio con nuevas y variadas tecnologías informáticas, y lograr la integración de soluciones de hardware, software existente.

Un primer estudio realizado, determinó la factibilidad de poder dar uso a microcontroladores (en este caso Intel Galileo 2.0) en la gestión de software y comunicación de resultados a través de una red de datos [1].

El presente trabajo, será impulsado como parte de sus actividades, por el responsable y adscriptos a la asignatura Redes de Datos del cuarto año de la carrera de Licenciatura en Sistemas de Información (LSI). El enfoque propuesto es dotar al Laboratorio de nuevos y variados servicios informáticos, buscando que las mismas se integren y aporten armonía para su correcto funcionamiento.

1. Introducción

Los aspectos a desarrollar en este trabajo son:

- Domótica. Definida por [2] como la integración de la tecnología en el diseño

inteligente de un recinto cerrado, o mejor aún como el uso simultáneo de electricidad, electrónica, informática y comunicaciones aplicadas a la gestión de las viviendas. Los componentes generales de un sistema domótico a estudiar son: el controlador, el sensor, el actuador y las interfaces.

- Red de datos. Se hace necesario para el despliegue, establecer mecanismos para lograr comunicación entre el controlador mencionado y las distintas aplicaciones de software, los cuales requieren estar enlazados por un determinado medio físico y utilizar un mismo protocolo de comunicación, para lugar a la denominada red de datos [3].

- Protocolos de red. En otro nivel, se requiere la aplicación de protocolos livianos con gestión de infraestructura, como Lightweight Directory Access Protocol o LDAP, conjunto de protocolos abiertos que se utilizan para acceder a la información almacenada de manera segura dentro de una red. Actualmente, LDAP se usa más dentro de organizaciones individuales, como universidades, departamentos gubernamentales y empresas privadas. El servidor de LDAP puede usar una variedad de bases de datos para guardar los directorios, cada uno optimizado para operaciones de lecturas rápidas y reiterativas. La principal ventaja de LDAP es que la información de toda una organización se puede consolidar en un repositorio central [4].

- Control y gestión de acceso. En el nivel de la aplicación informática, se debe garantizar que sólo las personas autorizadas tengan acceso a la información, que la información se mantenga intacta y disponible. El propósito es evitar el uso de la información de manera no autorizada, permitir el uso y modificación de la información por los usuarios autorizados, y preservar la consistencia interna y externa de los datos. La norma ISO 27002 describe varias áreas donde la gestión de acceso a los usuarios debe ser considerada: registro del usuario, gestión de privilegios, administración de contraseñas de usuario y revisión de los derechos de acceso de los usuarios [5].

- Control de acceso basado en roles (RBAC, Role Based Access Control). Técnica que simplifica la gestión de autorización y permisos en diferentes ambientes. Se desea dar reemplazo a sistemas de permisos de acceso que se conceden directamente al usuario, por sistemas de control de acceso basado en roles. Los roles permanecen estables en comparación a los usuarios. El rol de hecho es asociado con un conjunto de opciones de permisos en particular. Cuando los usuarios cambian, los roles solo necesitan ser retirados y reasignados [6].

- Autenticación en aplicaciones. Las aplicaciones web modernas presentan interrogantes a la hora de resolver el proceso de autenticación, al intentar hacerlo con métodos y herramientas convencionales. Las razones tienen que ver con las formas que actualmente se crean las aplicaciones y el entorno en el que se localizan y distribuyen. El despliegue de las aplicaciones actuales se hace sobre numerosos servidores, localizados en diferentes sitios por motivos de alta disponibilidad y distribución de cargas; se busca aumentar el tiempo de actividad y mitigar situaciones de alta latencia. El efecto secundario que surge es, cuando un usuario accede a una aplicación, ya no se garantiza que siempre esté accediendo al mismo servidor, dado que el usuario puede haber iniciado sesión en un servidor, pero no en los otros en los que se distribuye la aplicación. Queda en este sentido, resolver:

- Autenticación basada en cookies: ha sido el método predeterminado y comprobado para manejar la autenticación de usuarios durante mucho tiempo, la misma es “con estado”. Esto significa que un registro o sesión de autenticación debe mantenerse tanto en el servidor como en el lado del cliente. El servidor necesita realizar un seguimiento de las sesiones activas en una base de datos, mientras que en el front-end se crea una cookie que contiene un identificador de sesión [6].

- Autenticación basada en token: trata de un método “sin estado”, el servidor no mantiene registro de qué usuarios están conectados o de la emisión de token. En cambio, cada solicitud al servidor va acompañada de un token que el servidor utiliza para verificar la autenticidad [7].

- Comunicación entre aplicaciones de software. Las aplicaciones single-page, tienden a utilizar técnicas para recuperar y consumir datos JSON de una API REST “Transferencia de Estado Representacional”, refiere a un método de transferencia de servicios web que permite a diferentes equipos acceder y manipular representaciones textuales de recursos web mediante un set uniforme y predefinido de operaciones sin estado” [8]. Cuando los datos a comunicar son proporcionados por una API, presenta varias ventajas, una de ellas es que los datos se utilicen en más de una aplicación [9].

II. Líneas de investigación y desarrollo.

En la línea de Ingeniería de Software se proponen las siguientes actividades:

- Utilizar Microsoft Active Directory [10] como repositorio centralizado de usuarios con control de acceso basado en roles.
- Analizar los diferentes productos hardware para desarrollos de domótica.
- Estudiar el set de herramientas necesarias para el desarrollo de software.
- Desarrollar una API REST para la gestionar los actuadores.
- Utilizar JWT [11] para autenticar a los usuarios.
- Desarrollar la interfaz o front-end inicial de la aplicación móvil.
- Configurar una red WiFi para soportar la conexión de dispositivos de usuario.
- Establecer mecanismos de log de todos los procedimientos de comunicación de datos y las formas de monitoreo.

- Proponer formas de reporte de información acerca de las personas que acceden, tiempo de permanencia, entre otros aspectos.

III. Resultados.

Resultados obtenidos. El trabajo presentado en JAIIO 45 [1], determinó la factibilidad técnica de poder gestionar un sistema informático instalado en microcontroladores, y que, en base a mecanismos de comunicación de datos, fue posible la transmisión de información para su posterior gestión. Trabajos previos desarrollados en el curso de la asignatura del año 2017, determinaron:

- la factibilidad de utilizar Raspberry pi 3 [12] como controlador del sistema de domótica,
- el uso del lenguaje de programación Kotlin [13] para el desarrollo del front y back end.
- el uso del framework Spring Boot [14] para el despliegue de la API REST.

Bajo estas determinaciones, fue posible implementar la infraestructura, en un esquema como lo muestra la Figura 1, compuesta por:

- MS Active Directory: Gestión y control de usuarios
- Raspberry pi 3: Controlador del sistema de domótica
- API REST sin estado: Transferencia de servicios web y generación de tokens “JWT”
- Interfaz: aplicación front-end Android

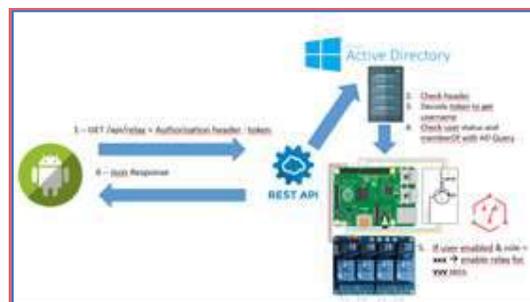


Figura 1. Esquema de funcionamiento.

Resultados esperados. Se está trabajando en la implementación en el laboratorio LRDTBD para dar continuidad a las pruebas.

- Se espera una vez implementada la solución, proponer a las autoridades de la Facultad su uso en modalidad Demo, para el análisis de funcionamiento.
- Los datos obtenidos por el monitoreo podrían ser utilizados por otras asignaturas en actividades de, por ejemplo, Data Mining, análisis estadístico, entre otros.
- Incorporar la funcionalidad de conocer el listado de personas que se encuentran en el edificio, o que hayan ingresado en algún momento determinado, como también su egreso. Si bien no está previsto incorporar el concepto de 'presencia', es decir, saber si una persona está o no en el edificio, es una funcionalidad que puede ser incorporada a futuro mediante por ejemplo la incorporación de tecnologías RFID (Identificación por radiofrecuencia)
- Conectar este sistema, con el servicio (en modalidad demo) de VoIP del Laboratorio LRDTBD, a efectos de poder, por ejemplo, contactar a las personas registradas en el edificio a través de una llamada telefónica VoIP (Voz sobre protocolo de internet).
- Incorporar cuentas de docentes y No-docentes de la Facultad, a la base de datos de usuario de Active Directory, para fortalecer y ampliar las pruebas de funcionamiento.

IV. Formación de recurso humano.

En el Grupo de Investigación en Innovación en Software y Sistemas Computacionales (GISSC) están involucrados 4 docentes investigadores, un becario de investigación de

pregrado, 1 tesista de doctorado y 3 tesistas de maestría.

Para el caso de esta línea de investigación en particular, se encuentra trabajando 1 docente investigador, y 2 Ayudantes Adscriptos (Licenciado en Sistemas de Información) a la asignatura Redes de Datos.

V. Referencias.

- [1] Trabajo en 45 JAIIO.
<http://45jaiio.sadio.org.ar/node/55>
- [2] R. Hernández Balibrea, Tecnología domótica para el control de una vivienda, Cartagena: Universidad Politécnica de Cartagena, 2012.
- [3] J. Kurose y R. Keith, Redes de Computadoras - Un enfoque descendente, vol. V, Pearson.
- [4] M. Heslin, Integrating Red Hat Enterprise Linux 6 with Active Directory, Red Hat Inc, 2013.
- [5] T. R. Peltier, «Information Security Fundamentals,» *Taylor & Frances Group*, n° 2a, 2014.
- [6] J. Yang-Feng, Z. Si-Yue, H. Zhen, L. Mu-Qing, Y. Ling y N. Jing-Ping, «Access control for rural medical and health collaborative working platform,» *The Journal of Chine Universities of Posts and Telecommunications*, n° 20, 2013.
- [7] Auth0, «<https://auth0.com/>,» Auth0, 31 Mayo 2016. [En línea]. Available: <https://auth0.com/blog/cookies-vs-tokens-definitive-guide/>.
- [8] C. Pautasso, E. Wilde y R. Alarcon, REST: Advanced Research Topics and Practical Applications, 2014.
- [9] L. Richardson y M. Amundsen, RESTful Web APIs, O'Reilly Media, 2013.
- [10] Microsoft, «Microsoft AD,» Microsoft, [En línea]. Available: <https://support.microsoft.com/es-es/help/196464>.
- [11] Auth0, «JSON Web Token,» Auth0, [En línea]. Available: <https://jwt.io/>.

- [12] Raspberry,
«<https://www.raspberrypi.org/>,»
[https://www.raspberrypi.org/products/
raspberry-pi-3-model-b/](https://www.raspberrypi.org/products/raspberry-pi-3-model-b/).
- [13] Jet Brains, «<https://kotlinlang.org/>,»
Jet Brains, [En línea]. Available:
<https://kotlinlang.org/>.
- [14] Pivotal, «spring,» Pivotal Software,
[En línea]. Available:
<https://projects.spring.io/spring-boot/>.

Tecnología aplicada al deporte de alto rendimiento

Laura Fava, Diego Vilches, Javier Díaz, Matías Pagano, Ramiro Romero Dapozo
LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.
Facultad de Informática. Universidad Nacional de La Plata, Calle 50 esq. 120, 2do Piso.
Tel: +54 221 4223528
{lfava, dvilches, jdiaz, matiasp}@info.unlp.edu.ar, rromero@linti.unlp.edu.ar

RESUMEN

Internet de las Cosas (IoT) es una nueva revolución de Internet. Los objetos se hacen reconocibles y obtienen inteligencia haciendo o permitiendo decisiones relacionadas con el contexto gracias al hecho de que pueden comunicar información sobre sí mismos, pueden acceder a información, o pueden ser componentes de servicios complejos. El término IoT se refiere generalmente a escenarios donde la capacidad de cómputo y la conectividad de las redes se extienden a objetos, sensores y elementos cotidianos permitiendo que estos dispositivos generen, intercambien y consuman datos (Karen R., 2015).

Por otro lado, en un mundo tan competitivo como es el del deporte, cualquier detalle por mínimo que sea, marca una gran diferencia; los datos que ofrecen los diferentes sensores aportan una ventaja competitiva extra que puede ser diferencial. El uso de *wearables*, la variedad de datos que pueden obtenerse a partir de ellos y la velocidad a la que se pueden procesarse dan paso a una nueva etapa donde los deportistas, entrenadores y técnicos, pueden aprovechar estas nuevas tecnologías para mejorar su habilidad, entrenamiento y resultados. Con sensores o sin sensores, la tecnología ha alcanzado la industria del deporte.

En este artículo se describen líneas de investigación y desarrollo que forman parte de un proyecto integral destinado a proveer soluciones tecnológicas que aporten a un entrenamiento más controlado, medido, estimulado y eficiente.

Palabras claves: Internet de las Cosas (IoT), sensores, *wearables*, protocolos para IoT, análisis de video.

CONTEXTO

El Laboratorio de Investigación de Nuevas Tecnologías Informáticas LINTI de la Facultad de Informática, viene trabajando en proyectos relacionados con Internet de las Cosas en áreas como domótica, horticultura, estacionamiento inteligentes, etc.

Uno de los primeros proyectos de IoT vinculado a ciudades inteligentes fue rParking: un sistema de plazas de estacionamiento reservadas (Boccalari, E., González, F., 2016). En rParking se analizó la problemática de las plazas de estacionamiento reservadas en un contexto de SmartCities y se propuso una solución abarcando las tecnologías correspondientes, tanto de software como de hardware. Este sistema se testeó y la red de sensores se encuentra en

funcionamiento en la playa de estacionamiento del edificio de Rectorado de la UNLP.

Otro proyecto con IoT es Electra, una red de sensores y una herramienta gráfica que visualiza el consumo de energía de los artefactos/dispositivos eléctricos de la Facultad de Informática de la UNLP. Esta red y su aplicación vinculada, están en una etapa final de desarrollo. La aplicación proveerá dos modalidades, una basada en web que facilitará el monitoreo del consumo de los dispositivos de las aulas de la facultad, pudiéndose administrar remotamente y una versión móvil, destinada a docentes, no docentes y alumnos de nuestra entidad, que permitirá visualizar en tiempo real el uso y consumo de energía, promoviendo un uso energético responsable y cuidado.

Asimismo se han propuesto soluciones de IoT para ayudar a personas con discapacidad como el calzado háptico, una solución tecnológica integral en forma de zapato que va acompañado por una aplicación móvil que guía a los disminuidos visuales a un destino específico en tiempo real (Berretti, F., 2014) y un desarrollo incipiente que consiste en un kit de bastón y anteojos con sensores que detectan y avisan ante objetos en espacios bajos y altos para personas ciegas.

Como puede observarse, se está trabajando en diferentes sublíneas dentro de IoT pero en este artículo nos referiremos especialmente a IoT en deporte.

Las tecnologías de IoT nos ofrecen una potente fuente de información para planear el progreso del deportista y optimizar sus entrenamientos y competiciones. Hoy en día, los profesionales del deporte pueden tener datos exactos y en tiempo real, con los

que medir su velocidad, la distancia que han recorrido, los movimientos realizados o su aceleración.

Las líneas de trabajo que se describen en este artículo se desarrollan en el LINTI y están enmarcadas en el proyecto *Internet del futuro: Ciudades digitales inclusivas, innovadoras y sustentables, IoT, ciberseguridad y espacios de aprendizaje del futuro*, acreditado en el marco del Programa de Incentivos, bajo la dirección del Lic. Javier Díaz.

1. INTRODUCCIÓN

El término Internet de las Cosas o Internet of Things (IoT) se refiere generalmente a escenarios donde la capacidad de cómputo y la conectividad de las redes se extienden a objetos, sensores y elementos cotidianos -no computadoras personales-, permitiendo que estos dispositivos generen, intercambien y consuman datos.

El concepto de combinar computadoras, sensores y redes para monitorear y controlar diferentes dispositivos ha existido durante décadas. Sin embargo, la reciente confluencia de diferentes tendencias del mercado tecnológico está permitiendo que la Internet de las Cosas esté cada vez más cerca de ser una realidad generalizada. Estas tendencias incluyen la conectividad omnipresente, la adopción generalizada de redes basadas en el protocolo IP, la economía en la capacidad de cómputo, los avances en el análisis de datos y el surgimiento de la computación en la nube.

Con todo esto, la implementación a gran escala de dispositivos de la IoT promete transformar muchos aspectos de la forma en que vivimos. Los nuevos productos de IoT para automatización del hogar y dispositivos de gestión de energía

nos están llevando hacia una visión de la “casa inteligente”, los vehículos conectados en red, los sistemas de tráfico inteligentes y los sensores integrados en calles, puentes o luminarias, nos acercan más a la idea de “ciudades inteligentes” y como era de esperar IoT también está alcanzando la industria del deporte.

La utilización de dispositivos inteligentes o wearables y el análisis de datos que generan en tiempo real están cambiando el mundo del deporte, y los deportistas podrían optimizar su rendimiento y adecuar sus entrenamientos utilizando la tecnología IoT.

En este artículo se describen líneas de investigación y desarrollo que forman parte de un proyecto integral destinado a proveer soluciones de IoT vinculadas con el entrenamientos de deportistas de élite y destinadas a mejorar su entrenamiento y rendimiento.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Las líneas de investigación, desarrollo e innovación que se llevan a cabo en este proyecto están vinculadas al desarrollo de dispositivos basados en sensores y al desarrollo de aplicaciones que permitan mejorar el entrenamiento de los deportistas de alto rendimiento.

Los ejes principales de I+D+i son:

- Análisis de tecnologías de vanguardia para la construcción de dispositivos y software que mejore el entrenamiento y rendimiento de los jugadores.
- Diseño y construcción de dispositivos basados en LEDs (light emitting diodes), sensores de proximidad y microcontroladores para mejorar la toma de decisiones, la reacción física y cognitiva, la capacidad de atención y

concentración de los jugadores.

- Diseño y construcción de dispositivos basados en GPS (Global Positioning System) y acelerómetro, para medir parámetros como máximas velocidades, impactos, distancia recorridas, etc. en situación de juego.
- Diseño e implementación de una aplicación móvil que permita el registro de eventos de un partido en tiempo real para ser analizados en el entretiempo y luego de finalizado el mismo (Kröckel, P., Piazza, A. & Neuhofer, K., 2017) . Asimismo se trabajará con los datos registrados para ser exportados a diferentes plataformas de análisis de video como LongoMatch (LongoMatch) o Wyscout (Wyscout).

3. RESULTADOS Y OBJETIVOS

Para esta línea de IoT y deporte se han analizado diferentes kits de desarrollo para sistemas embebidos, sensores, protocolos de comunicación, así como también los beneficios que implicaría la creación de dispositivos para deporte.

Como resultado se han creado prototipos de hardware, que consisten en dispositivos (o tortuguitas) compuestos por un módulo Wi-Fi y un conjunto de LEDs RGB que pueden ser configurados y operados desde dispositivos móviles para crear las estrategias del entrenamiento y analizar el rendimiento de los deportistas mediante un prototipo de software que también se ha implementado. También se diseñó e implementó un protocolo de red binario personalizado para el control de las

tortuguitas desde un dispositivo móvil Android. La Fig. 1 muestra un primer diseño de los dispositivos y la placa de circuito impreso (PCB) que sobre la cual se montan los circuitos integrados que constituyen el dispositivo.

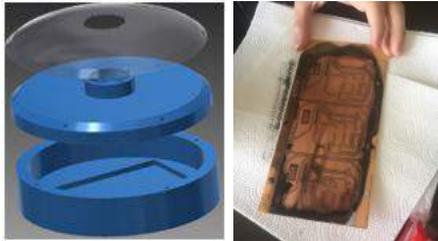


Fig 1.: Diseño de dispositivos basados en LEDs

Este primer prototipo pudo ser probado en entrenamientos de un club amateur y en un club profesional. La Fig. 2 muestra un escenas de tales pruebas. Las pruebas permitieron testear el comportamiento de los dispositivos en exterior y en interior y encontrar nuevas funcionalidades para la próxima versión del software de administración.



Fig 2.: Pruebas de prototipos

También se ha trabajado para registrar información en situación de juego usando el sistema de posicionamiento global o GPS. Para esto se ha diseñado y creado un prototipo de dispositivo con un módulo receptor de posicionamiento global (GPS) y para tener más precisión se ha adicionado un acelerómetro. La Fig. 3. muestra el primer prototipo de

hardware del dispositivo, en cuanto al software se han realizado pequeñas pruebas de conectividad y recepción de información pero no se ha alcanzado a implementar un prototipo de software.

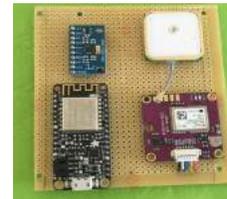


Fig 3.: Prototipo con GPS y acelerómetro

Los objetivos inmediatos son:

- Re-diseño y construcción de una nueva versión del dispositivo basado en LEDs considerando nuevos sensores, redistribución de las piezas y una carcasa más adaptable a los diferentes deportes (conos, palos, pared, etc.).
- Evaluación de protocolos estándares de comunicación como MQTT y COAP para determinar si se reemplazan por el protocolo creado.
- Re-diseño y construcción de una nueva versión del dispositivo basado en GPS considerando los nuevos receptores GPS.
- Implementación de una aplicación que registre los datos en tiempo real de los GPS y produzca información útil como máximas velocidades, distancias recorridas, gravedad de impactos, etc.
- Diseño y creación de una aplicación móvil que permita el registro de eventos de un partido en tiempo real para ser analizados en el entretiempp y luego de finalizado el mismo. Exportación de datos de eventos a diferentes plataformas de análisis de video.
- Integración de datos generados por las diferentes fuentes con las que

trabajamos, GPS, acelerómetro, análisis de video, fichas deportivas, médicas, plataformas de saltos, etc.

- Generación de información útil en tiempo real no sólo para planificar los entrenamientos sino para la toma de decisiones durante la competencia (Zeng, J. & Jia J., 2017).

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de la línea de I+D+i presentada en este artículo se encuentra formado por docentes investigadores categorizados del LINTI y alumnos avanzados de la Licenciatura en Informática, Licenciatura en Sistemas e Ingeniería en Computación perteneciente a Facultad de Informática y a la Facultad de Ingeniería.

En relación a las tesis de grado vinculadas con esta línea de investigación, se está dirigiendo a 3 tesis. Se han realizado 3 Prácticas Profesional Supervisadas (PPS) y se encuentran 4 más en desarrollo. También se están ejecutando dos proyectos acreditados de la Secretaría de Políticas Universitarias del Ministerio de Educación y Deportes de la Nación.

5. REFERENCIAS

Berretti, F. *Calzado Háptico: Navegabilidad Asistida para Personas con Disminución Visual*. Tesis de Grado accesible en <http://sedici.unlp.edu.ar/handle/10915/63352>. Diciembre 2014.

Karen R., Scott E., Lyman C., *La Internet de las cosas, una breve reseña*. Internet Society (ISOC), Octubre 2015.

Kröckel, P., Piazza, A. and Neuhofer, K. *Dynamic Network Analysis of the Euro 2016 Final: Preliminary Results*, 2017 5th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), Prague, 2017, pp. 114-119.

Zeng, J. and Jia J., *The impact of big data on school sports and competitive sports*, 2017 Chinese Automation Congress (CAC), Jinan, 2017, pp. 596-599.

RECURSOS/DESARROLLOS EXISTENTES

LongoMatch: análisis de video, accesible en <https://longomatch.com/es/pro/>, 2018.

Wyscout: professional football platform for football analysis, accesible en <https://wyscout.com/>, 2018.

GPSports: The athlete tracking system, accesible en <http://gpsports.com/>

Catapult: the tracking system, accesible en <https://www.catapultsports.com/>

Data-driven Agriculture, las tecnologías aplicadas a problemas de agricultura regional

Alberto Eduardo Riba⁽¹⁾, Fernanda Beatriz Carmona⁽¹⁾, Jorge Damián Tejada⁽¹⁾, Matías Agustín Pérez⁽¹⁾, Emmanuel Alejandro Portugal Murcia⁽¹⁾, Nelson Acosta^(2,3), Juan Manuel Toloza^(2,3)

¹ Departamento de Básicas y Aplicadas, Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina
{ariba, jtejada, fbcarmona}@undec.edu.ar
{emmanuel.portugal.91, mataguper}@gmail.com

² Universidad Nacional del Centro de la Provincia de Buenos Aires
General Pinto 399, Tandil, Buenos Aires, Argentina
{nacosta, jmtoloza}@exa.unicen.edu.ar

³ Universidad Nacional de Tres de Febrero
Mosconi 2736 - Sáenz Peña (B1674AHF), Buenos Aires, Argentina

Resumen

El sector agrícola regional se esfuerza por aplicar prácticas agrícolas modernas que garanticen una provisión segura y sostenible de alimentos de calidad, fomentar la eficiencia del uso de recursos, desarrollar una economía rentable y combatir el cambio climático.

La agricultura de precisión cumple un rol muy importante en estas prácticas, los avances en herramientas de sensado, la generalización del uso de los sistemas de posicionamiento globales, el uso de robots e Internet de las cosas (IoT) permiten mejorar la productividad y optimizar el uso de recursos agrícolas mediante el uso intensivo de los datos generados.

Data-Driven Agriculture es una nueva tendencia que no solo implica la digitalización y almacenamiento de información sino, también mediante el uso de herramientas de Big Data, desplegar políticas de gestión de los datos enfocadas a su análisis, explotación y protección.

Esta línea de I+D+i de gran interés regional, se enfoca en el problema de la racionalización y uso eficiente de los recursos agrícolas en regiones de climas áridos, con el objetivo de incrementar la productividad y combatir el cambio climático. Problemas detectados en empresas privadas y organizaciones del medio.

Los temas abordados son transversales a las áreas, programación, probabilidad y estadística, investigación operativa, bases de datos y análisis de sistemas.

En los proyectos enmarcados en esta línea participan alumnos de grado y los desarrollos propuestos convergen en tesinas de la Ingeniería en Sistemas y Licenciatura en Sistemas de la UNDeC.

Palabras clave:

Agromática, Data-Driven Agriculture, Red de Sensores, Agricultura de Precisión, Sistemas GNSS, Series de Datos Agrometeorológicos.

Contexto

Esta línea de I+D+i corresponde al desarrollo e implementación de proyectos que fortalecen la inserción de la UNdeC en la comunidad y especialmente en el medio productivo de la región y refiere a los proyectos: “Programación y generación de pronósticos de riego presurizado para cultivo de olivo en regiones áridas, utilizando TIC” FICyT - UNdeC 2009-2011, “Integración y administración de índices de sensado en la programación y pronósticos de riego presurizado para cultivo de olivo en regiones áridas”, FICyT - UNdeC 2012 - 2014, “Utilización de métodos de diseño de software para desarrollar un sistema automatizado de riego”, FICyT - UNdeC 2011 - 2013, “Red de Sensores Inalámbricos basado en microcontroladores para la monitorización del riego presurizado en plantaciones de olivo”, FICyT - UNdeC 2013-2015, “Incremento de la precisión posicional relativa utilizando receptores GPS de bajo costo” FICyT - UNdeC 2016-2018, aprobados por la Secretaría de Ciencia y Tecnología, convocatoria para estímulo y desarrollo de la investigación científica y tecnológica (FICyT - UNdeC).

Además, se coopera activamente con distintos grupo de trabajo dentro de la universidad que desarrollan trabajos relacionados a la línea y con una empresa del medio dedicada al cultivo de olivo, esperando que otras empresas de la región se beneficien de los resultados de esta línea de I+D.

Introducción

La actividad agrícola en la región sólo es posible con la ayuda del riego artificial, utilizándose las tierras para cultivos de nogal, olivo, vid y frutales. La

agronomía es la ciencia aplicada que rige las prácticas agrícolas y es considerada una ciencia espacial. En un lote cultivado es posible encontrar sectores de alta productividad, muy próximos a sectores menos productivos. Esta variación espacial suele estar asociada a factores como la pendiente del suelo, la permeabilidad, el tipo del suelo y la fertilidad. Sin embargo, los agricultores manejan el cultivo de forma homogénea, aplicando dosis de fertilizante o irrigando de manera uniforme todo el lote.

La agricultura de precisión utiliza complejas fórmulas y modelos matemáticos para el análisis de los grandes volúmenes de datos geo-espaciales generados por las distintas tecnologías de sensado, convirtiéndose en un sistema de control en donde la retroalimentación de la información permite a los productores diseñar tratamientos específicos situados para incrementar la eficiencia en el uso de los recursos involucrados [1].

El sector agrícola es el mayor consumidor de agua, por el aumento de la superficie irrigada y por la escasa eficiencia en los sistemas de riego. Por este motivo es de vital importancia lograr un equilibrio hidrológico que asegure el abastecimiento de agua a la población y al sector agrícola.

En regiones de climas áridos el costo final de explotación es afectado en un alto porcentaje por factores imputables al riego. Dentro de este costo se considera la inversión inicial del sistema (detección de las napas, perforación del pozo y tendido del sistema de distribución) y el costo energético para su extracción y distribución (energía eléctrica y combustibles necesarios).

El sistema de irrigación más utilizado por los agricultores es el riego presurizado. Desde el punto de vista agronómico se denominan riegos

localizados porque humedecen un sector de volumen de suelo suficiente para un buen desarrollo del cultivo.

Para incrementar la eficiencia en los sistemas de riego es necesario estudiar las distintas variables (requerimientos hídricos de los cultivos, características del suelo, condiciones meteorológicas, propiedades y limitaciones del sistema de riego) y como se relacionan para determinar el uso adecuado del recurso. Estas variables son de naturaleza heterogénea y algunas de ellas pueden ser capturadas automáticamente mediante el uso de sensores para permitir su posterior tratamiento y análisis para una correcta programación del riego.

Monitorización de riego en cultivos

Para la monitorización y programación del riego se debe tener en cuenta:

Factores ambientales:

- la medición o estimación de variables ambientales: temperatura, presión, radiación solar, evapotranspiración, entre otras.
- el monitoreo en el crecimiento y desarrollo del cultivo o fitomonitorio: tamaño del fruto, tamaño del tallo, flujo de savia, índice de estrés hídrico.
- la medición de variables en el suelo: temperatura, humedad, conductividad, constante dieléctrica.
- la medición de variables propias del agua aportada: salinidad, alcalinidad, etc.

Factores artificiales:

- Limitantes propios del sistema de riego instalado (cañerías, presiones, válvulas, aspersores, goteros). Entre ellos perfil-umbral (capacidad máx. diaria), coeficiente de uniformidad, etc.
- Aspectos económicos relativos al consumo energético utilizado por las bombas. Los aranceles de Kwatt/hora

varían según el momento del día en que son consumidos.

Actualmente las parcelas de la región capturan la información relacionadas con algunas de estas variables de manera aislada e independiente por tipo de sensor (de temperatura ambiente, de humedad de suelo, dendrómetros, estaciones meteorológicas, dataloggers, etc.), utilizando para el análisis e interpretación de datos la interface de software suministrada por el fabricante del dispositivo. Para otras variables el método de registro es totalmente manual utilizando planillas en papel y dicha información luego es volcada en planillas de cálculo.

Como resultados de proyectos de esta línea de investigación se implementó un sistema de registro para el procesamiento de esta información [2, 3, 4]. Como continuación se desarrolló una estación inalámbrica que automatiza la captura de la información de sensores ubicados en distintos sectores de la plantación [5]. Cada nodo de la red está compuesto por un dispositivo inalámbrico autónomo y un conjunto de sensores para la recolección de datos de naturaleza agrometeorológica. Esta práctica ha sido implementada con éxito en diversos ámbitos como detección de incendios forestales [6] o la monitorización de viñedos [7]. Estas redes se caracterizan por su escalabilidad, ausencia de cableado y bajo consumo, lo que las vuelve muy interesantes para aplicaciones en agricultura, ya que serían más costosas y complejas de implementar con otras tecnologías.

A partir de la automatización de la recolección de información de los sensores y la integración con el sistema actual, se optimiza la generación de estimaciones estadísticas y gráficos, lo que contribuirá al proceso de toma de decisiones relacionadas con el pronóstico y la planificación del riego.

Riego automatizado inteligente

Existe una gran cantidad de sistemas para la determinación, control y automatización del riego que permiten un consumo óptimo de agua.

En nuestra región, los equipos de riego localizado empleados en la mayoría de las explotaciones agrícolas no poseen automatismos, o en algunos casos, estos se encuentran en desuso. La apertura y cierre de las válvulas de las diferentes subunidades y sectores de riego se realiza en forma manual, acotando los tiempos de riego, basándose principalmente en la experiencia de quien lo programa o en recomendaciones surgidas de mediciones de humedad de suelo. La fertirrigación sufre de las mismas limitaciones.

Ajuste de series meteorológicas

La automatización de la recolección de información de sensores permitió notar que las series agrometeorológicas resultantes presentan problemas de completitud, veracidad y exactitud.

La calidad de estas series depende de dispositivos electromecánicos (sensores, estaciones meteorológicas, dataloggers), de redes de transmisiones inalámbricas, de personal de mantenimiento y de baterías. Cualquiera de estos factores resulta un punto de fallo ineludible que degrada la calidad [8]. Para mitigar estos se pueden establecer políticas y estándares de funcionamiento de hardware y software que tiendan a prevenirlos, minimizarlos y ante su aparición, controlarlos. A pesar de la implementación de tales políticas existen casos en los que simplemente los fallos ocurren, y se deben tomar medidas ante estas situaciones.

Todos estos motivos inclinaron al estudio, análisis y desarrollo de métodos y técnicas de detección, corrección y ajuste de datos de series agrometeorológicas para mejorar su calidad. Actualmente un trabajo final de

grado de un alumno avanzado de la carrera está relacionado con esta temática.

Los Sistemas Globales de Posicionamiento (GNSS)

Los sistemas GNSS (Global Navigation Satellite System) son muy utilizados en la agricultura de precisión para georreferenciar sectores de parcelas, existen tecnologías implementadas como NAVSTAR-GPS y GLONASS, y otras en vías de implementación como GALILEO y COMPASS.

En la actualidad existe una gran variedad de receptores GNSS y su costo varía en función de su precisión. Los receptores de bajo costo proveen posiciones con precisiones de una decena de metros (aplicables a navegación, seguimiento de fauna y flotas, turismo y ocio, etc.). Cuando se trabaja con este tipo de receptores no se puede asegurar de obtener posiciones con una precisión mayor a los 15 metros el 95% de las veces [9, 10].

La agricultura de precisión necesita de sistemas que provean más precisión en sus localizaciones para optimizar el uso de los recursos, siendo normalmente estos dispositivos más costosos [11]. Por este motivo es muy importante estudiar, diseñar y desarrollar de algoritmos, técnicas y métodos que permitan disminuir el error en la posición entregada por receptores GNSS de bajo costo para mejorar la precisión del posicionamiento. Esta temática es abordada en una tesis de maestría por uno de los integrantes del equipo de trabajo.

Líneas de Investigación, Desarrollo e Innovación

- Sistemas de Información y Base de Datos
- Sistemas de Tiempo Real

- Diseño de dispositivos de automatización
- Comunicaciones inalámbricas
- Redes de sensores
- Ajuste de series de datos
- Posicionamiento de precisión
- Protocolo NMEA
- Sistemas Multiconstelación
- Posicionamiento diferencial de bajo costo

Resultados y Objetivos

Objetivos

- Analizar y estudiar las diferentes tecnologías de microcontroladores, módulos de conexiones inalámbricas y tipos de sensores (de temperatura ambiente, de humedad de suelo, dendrómetros) disponibles en el mercado.
- Monitorizar en forma centralizada, remota y en tiempo real las variables capturadas por los diferentes sensores.
- Mejorar la capacidad de administración y planificación de los recursos hídricos destinados al riego, a través del análisis de las variables obtenidas de los nodos instalados en sectores con diferentes características de suelo y clima.
- Configurar adecuadamente los algoritmos de control y aplicar estrategias de riego que optimicen la relación kg. producido por m³ de agua aplicada.
- Evaluar distintos esquemas de control, comparar los resultados e inferir en la elaboración de nuevas estrategias de riego.
- Desarrollar técnicas, métodos y algoritmos para mejorar a la precisión del posicionamiento utilizando receptores GNSS de bajo costo en un prototipo de GNSS diferencial.

- Aumentar la uniformidad y la eficiencia de la aplicación del agua del riego para reducir los costos asociados de consumo del agua.

Resultados

Como resultados se puede mencionar; los trabajos de tesis de grado de dos de los integrantes, denominados “Sistema de Gestión de Riego y Fertilización”, presentado en 2014 y “Estación inalámbrica basada en microcontroladores para la monitorización del riego en plantaciones de olivo” presentado en marzo de 2018; el artículo presentado en el evento 45 JAIIO – Concurso de trabajos Estudiantiles 2016, titulado “Red de sensores inalámbricos basados en microcontroladores para la monitorización del riego presurizado en plantaciones de olivo”; el artículo presentado en XXII CACIC en 2016 titulado “Procesamiento de sentencias NMEA-0183 para el análisis de la geometría satelital utilizando receptores GPS de bajo costo”; el artículo presentado en XXIII CACIC en 2017 titulado “Estación de monitoreo en tiempo real de parámetros agrometeorológicos para determinar la necesidad de riego en plantaciones agrícolas”.

Formación de Recursos Humanos

El equipo de trabajo está formado por cinco docentes de las carreras Ingeniería en Sistemas, Licenciatura en Sistemas e Ingeniería Agronómica de la UNdeC (acreditadas por CONEAU), de la carrera Ing. Sistemas de la UNICEN y UNTREF y dos alumnos de la UNdeC.

De los docentes: 1 es posdoctorado en Informática, 1 es doctor en Informática; 2 maestrandos que presentarán su tesis en la Universidad Nacional de San Juan; un doctorando que presentará su tesis en la

Universidad Nacional del Centro de la Provincia de Buenos Aires.

Los alumnos de grado se hallan realizando su trabajo de tesina final en esta línea de I+D.

Los integrantes son docentes de las asignaturas Arquitecturas de computadoras II, Programación I y II y Agromática I y II. Estas asignaturas fomentan la participación en proyectos de investigación, por lo que pueden surgir nuevos trabajos en esta línea.

Referencias

- [1] E. W. Schuster, S. Kumar, S. E. Sarma, J. L. Willers and G. A. Milliken (2011) "Infrastructure for data-driven agriculture", IEEE 8th International Conference & Expo on Emerging Technologies for a Smarter World, New York, NY
- [2] F. B. Carmona, A. Riba, A. Sfeir, and F. E. Frati, "Pronósticos de riego presurizado para cultivo de olivo en regiones áridas," (2008), eje: Ingeniería de Software y Base de Datos.
- [3] F. B. Carmona, E. A. Riba, A. Sfeir, and F. E. Frati, "Pronósticos de riego presurizado para cultivo de olivo en regiones áridas," (2010), eje: Innovación en Sistemas de Software. [Online]. Available:<http://hdl.handle.net/10915/19571>
- [4] E. A. Riba, F. B. Carmona, F. E. Frati, J. D. Tejada, N. Acosta, and J. M. Toloza, "Integración y administración de índices de sensado en la programación y pronósticos de riego presurizado para cultivo de olivo en regiones áridas," (2012), eje: Innovación en sistemas de software.
- [5] C.-Y. Chong and S. Kumar, "Sensor networks: evolution, opportunities, and challenges," Proceedings of the IEEE, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.
- [6] Javier Solebera, "Detecting forest fires using wireless sensor networks," Sep. 2010.
- [7] Alberto Bielsa, "Smart agriculture Project in Galicia to monitor a vineyard with waspmote," Jun. 2012.
- [8] M. N. Khaliq, T. B. M. J. Ouarda (2007) "Short communication on the critical values of the standard normal homogeneity test (SNHT)" International Journal of Climatology, Vol. 27 681-687p
- [9] Tolosa Juan Manuel. (2012) "Algoritmos y técnicas de tiempo real para el incremento de la precisión posicional relativa usando receptores GPS estándar". SEDICI, Universidad Nacional de La Plata.
- [10] Alberto Riba, (2016) "Procesamiento de sentencias NMEA-0183 para el análisis de la geometría satelital utilizando receptores GPS de bajo costo". XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016).
- [11] Alberto Eduardo Riba, Jorge Damián Tejada, Fernando Emmanuel Frati, Nelson Acosta, Juan Manuel Toloza "Aumento de la precisión posicional empleando técnicas y algoritmos para el tratamiento del error en receptores GNSS de bajo costo", XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017)

Técnicas y estrategias para evaluar las ideas en la plataforma IdTR

Sandra Oviedo, Daniel Díaz, Alejandra Otazu, Francisco Ibañez,
Juan M. Cúneo, Elisa Silvia Oliva

Laboratorio de Tecnologías Informáticas Aplicadas a la Industria e Innovación
Instituto de Informática / Dpto. de Informática/ FCEfYn / UNSJ
CUIM – Av. Ignacio de la Roza 590 (O), Rivadavia – J5402DCS San Juan, 0264 4265101
{ddiaz, soviedo, aleotazu, fibanez}@iinfo.unsj.edu.ar; jmacuneo@gmail.com

Resumen

En un presente donde las tecnologías confluyen ofreciendo posibilidades de desarrollo de productos tecnológicos de manera casi ilimitada, toman relevancia las ideas porque ellas marcan la diferencia competitiva en diversos ámbitos. Y aunque el proceso de obtener nuevas ideas parece misterioso, la generación de ideas puede ser visto como un proceso relativamente estructurado y explicable, esta es la propuesta de la plataforma tecnológica IdTR -Ideas en tiempo real. La evaluación y selección de las ideas es una tarea compleja pues determina qué idea de una sesión de creatividad tiene más potencial innovador que otra. Una mala selección puede devenir en el descarte de una idea fructífera en el momento de su nacimiento. En IdTR no existe un proceso ni herramienta que asista a la evaluación y selección de ideas, este proceso es llevado a cabo subjetivamente por los sujetos que utilizan IdTR. Este proyecto propone desarrollar un proceso formal que permita asistir a la evaluación y selección de ideas susceptible de ser incorporado a la plataforma IdTR.

Palabras clave: Generación de Ideas- Programación Web en Tiempo Real- Evaluación de Ideas- Selección de Ideas

Contexto

Este trabajo describe los progresos en la investigación y estrategias que se están llevando a cabo encuadrados en el proyecto “IdTRES. Evaluación y Selección de Ideas en la Plataforma Ideas en Tiempo Real”, que se desarrolla en el Laboratorio de Informática Aplicada a la Innovación del Instituto de informática de la Facultad de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de San Juan.

Introducción

Quando se habla de innovación se hace referencia ya sea a un producto, un proceso, un servicio o a un modelo de negocio total o parcialmente novedoso puesto con éxito en el mercado, es decir, aceptado y consumido por el gran público, a diferencia de un invento, que aun tratándose de una novedad no llega al mercado. Se trata de un concepto de la Ingeniería Industrial [1].

El proceso que lleva a una innovación, llamado proceso innovador, se inicia con la instancia de creatividad, donde se concibe inicialmente una idea y luego mediante sucesivos estudios, y aplicación de diferentes herramientas y técnicas, esa idea inicial se mejora hasta lograr la concepción definitiva de la misma, es cuando se pasa a la puesta en producción del producto concebido y es llevado al mercado [2]. En trabajos precedentes, el equipo que hace esta propuesta obtuvo un prototipo de la Plataforma Ideas en

Tiempo Real (IdTR), se trata de un conjunto de herramientas que asisten al proceso de generación de ideas [3]. El cual consiste en: dada una temática, alrededor de la misma se generan ideas. De todas las ideas obtenidas, se elige la mejor para pasar a las instancias de prototipado y en el mejor de los casos de implementación. Uno de los problemas que presenta este proceso de generación de ideas esta en la selección de ideas, elegir la idea es un proceso complejo. En la versión actual de la plataforma la selección de ideas se deja en manos de los participantes, son ellos los que seleccionan la idea que prototiparan, no existe ninguna metodología ni herramienta que ayude en este proceso. Así planteado, este proceso puede conducir a una mala toma de decisión pues no se dispone de información, ni método que ayude a la toma de decisión, por lo tanto la selección de la idea queda librada al juicio subjetivo del sujeto que selecciona la idea.

Para elegir una idea, se pueden aplicar diferentes criterios y cada uno con sus argumentos podrán decir que una idea es mejor que otra si se considera más original, o técnicamente más factible, o económicamente más viable, o de mayor interés social, etc. Incluso, es probable que aunque una idea sea la mejor para algunos no lo sea para otros. Para el tratamiento y análisis de este tipo de problemas se han desarrollado métodos denominados de Decisión Multicriterio o Multiobjetivo (MDM). Desarrollar una herramienta de evaluación de ideas, implicará estudiar e implementar técnicas de análisis multicriterios, entre otros temas. Una aplicación así sería muy útil, ya que daría completitud a la gestión de ideas de la plataforma Ideas en Tiempo Real.

Líneas de Investigación

Lenguajes, técnicas y herramientas para la programación web de tiempo real

Existen hoy en día una buena cantidad de herramientas, frameworks y ambientes de desarrollo que se desarrollaron con el objeto de construir sistema web de tiempo y otros que han sido adaptados para soportar la programación web de tiempo real. En [4] se describen algunas tecnologías y plataformas utilizadas para desarrollar aplicaciones de tiempo real. Entre ellas se pueden destacar las plataformas web de tiempo real que son full-stack, y dentro de tales plataformas se va a describir Meteor [5] por ser la elegida para el desarrollo de ideas en tiempo real. Meteor es un framework desarrollado en javascript que utiliza el tiempo real para que las interacciones entre el cliente y el servidor tengan lugar a una velocidad tal que los usuarios experimentan retroalimentación directa sobre sus interacciones con la aplicación y otros usuarios [6]. Los datos modificados se muestran directamente sin que el usuario tenga que actualizar la página o el sistema deba ejecutar verificaciones periódicas. Meteor también suscribe al paradigma de programación reactiva [7], este permite construir elementos reactivos los cuales son seguidos internamente por Meteor y los cambios que se producen en estos son propagados hacia sus dependencias. Las tecnologías sobre las cuales está construido Meteor son Node.JS y MongoDB. El lenguaje de programación es JavaScript, tanto en servidor como en el cliente, el formato de intercambio de datos es JavaScript Object Notation (JSON). Una característica importante de Meteor es la capacidad de gestionar

bibliotecas de paquetes, existen dos grandes repositorios de paquetes para Meteor, [8], [9] el primero es un repositorio de paquetes nativos desarrollados para Meteor, mientras que el segundo es el repositorio global para los paquetes JavaScript. Este último fue introducido a partir de la versión 1.3 y potenció notablemente las capacidades de desarrollo del framework. Otra característica importante es que el servidor envía datos, no HTML, y el cliente los procesa.

La Plataforma Ideas en Tiempo Real

Cuando se habla de herramientas de software que asisten a la creatividad y obtención de ideas en general, se encuentran varias de diferente naturaleza, para diferentes propósitos. Se pueden encontrar herramientas para asistir al proceso creativo facilitando el trabajo colaborativo, otras que facilitan la tarea de los diseñadores y artistas y otras que permiten editar ideas en sesiones de brainstorming [10], [11].

La mayoría de las herramientas disponibles en la web dan soporte a la técnica de creatividad conocida como brainstorming, lluvia de ideas, sin embargo, para la obtención de ideas para productos tecnológicos es necesario que este tipo de técnicas (brainstorming) esten insertas en un proceso de generación de ideas. Este proceso incluye una serie de pasos que permiten obtener ideas, el armado de grupos, la administración de las ideas permitiendo que estas sean resultado de la colaboración y otras habilidades para innovación, tales como, creatividad, gestión, trabajo colaborativo, pensamiento crítico, además de las habilidades técnicas [2] La implementación de Ideas en Tiempo Real (IdTR) se lleva a cabo con Meteor.js y su arquitectura de software está ligada las características que presenta Meteor.js. En

[3] se presenta la arquitectura del sistema, esta se divide en dos grandes partes, el cliente y el servidor. El cliente se ejecuta en un browser. El servidor se ejecuta bajo Node.js, éste es remoto y se ejecutará en algún proveedor de la nube. En el cliente, hay 3 tipos de módulos, los de la interface de usuario, los de datos reactivos, y los de compensación de latencia.

Métodos de Decisión Multicriterio

Los modelos multicriterio, son más flexibles que los unicriterios y se pliegan con mayor fidelidad a la demanda y a la práctica de los decisores y analistas. Existen algunas metodologías que se insertan en contextos muy diversos para la toma de decisiones. Brindan las mejores alternativas decisorias con objeto de proponer, si no las soluciones óptimas, por lo menos las "mejores" soluciones de compromiso para los problemas bajo análisis [12].

Existen varios métodos para abordar la toma de decisiones multicriterio, el antecedente inmediato de estos métodos es el denominado Cuadro de Puntuación, procedimiento mediante el cual se descompone un objetivo complejo, de múltiples dimensiones, en sus atributos constitutivos más relevantes. Cada uno de estos atributos recibe una ponderación relativa que mide su importancia en la consecución del Objetivo General (Objetivo Complejo) con lo que -a la vez- se establecen valoraciones relativas entre los atributos que conforman el conjunto. En relación con estos atributos se establecen criterios de satisfacción que son valorados por medio de escalas, que pueden ser de distinta naturaleza, tales como escalas cardinales, ordinales, nominales, dicotómicas. Precisamente, esta diversidad de escalas es la que permite capturar el grado de satisfacción de criterios sustantivamente diferentes, sean cuantitativos o cualitativos.

Las debilidades principales del Cuadro de Puntuación residen: 1) la asignación de peso relativo a los criterios se hace a partir de la subjetividad del analista, aunque esta pueda ser atenuada mediante el uso de ciertos procedimientos como el método Delphi o la ronda de expertos; y 2) En la posibilidad de que los procedimientos de puntuación contengan inconsistencias que no pueden ser verificadas.

Estas limitaciones pueden ser solucionadas –en buena medida– mediante la utilización de estos MDM, pues con ellos se puede reducir las afirmaciones conjeturales, las conclusiones no explicitadas o el comportamiento intuitivo que está íntimamente asociado al proceso de atribución de pesos relativos a los criterios y a la valoración de la medida en que las distintas soluciones propuestas (alternativas) satisfacen los objetivos específicos [13].

1. AHP (analytic hierarchy process). Este método propuesto por Thomas Saaty consiste en 4 etapas e involucra un grado medio de complejidad matemática, [14]:

- Definir el problema y determinar el área de conocimiento en que se localiza.
- Estructurar una jerarquía de decisiones con el objetivo principal en un nivel más alto.
- Construir una matriz de comparación de conjuntos de pares, cada elemento de un nivel superior se compara con el inmediato que le sigue.
- Usar las prioridades obtenidas como pesos en el nivel inmediatamente siguiente (para cada elemento). Se realiza de manera iterativa, hasta que se tienen todas las prioridades de las alternativas, hasta el nivel más bajo.

2. TOPSIS (Technique for Order Preference by Similarity to Ideal Solution). La técnica TOPSIS se basa en

el concepto que es deseable que una alternativa determinada se ubique a la distancia mas corta respecto de una solución ideal positiva y a la mayor distancia respecto a una solución ideal negativa. Una solución ideal se define como un conjunto de niveles (o puntuaciones) ideales respecto a todos los atributos considerados de un determinado problema, aun cuando la solución ideal usualmente sea imposible o no sea factible de obtener. En consecuencia, desde este punto de vista la racionalidad de la conducta humana consiste en ubicarse lo más cerca posible de tal solución ideal y en alejarse lo más posible de una solución anti ideal o ideal negativa. TOPSIS define un índice llamado similaridad (o proximidad relativa) respecto a la solución ideal positiva combinando la proximidad a la solución ideal positiva y la lejanía respecto a la solución ideal negativa. Se selecciona aquella alternativa que se ubica lo más cerca posible a la máxima similaridad respecto a la solución ideal positiva [15].

3 DEA (Data Envelopment Analysis). El Análisis Envoltente de Datos o Data Envelopment Analysis (DEA) es una poderosa técnica de optimización, desarrollada por Charnes, Cooper y Rhodes, construida para medir el comportamiento relativo de diferentes unidades organizacionales en las cuales la presencia de múltiples insumos (inputs) y productos (outputs) hace difícil la comparación de su desempeño. DEA permite comparar la gestión relativa de un grupo de unidades de producción de bienes y/o servicios que utilizan el mismo tipo de recursos (insumos) para producir un mismo grupo de productos (salidas). La metodología identifica fronteras eficientes y permite hallar indicadores de gestión relativa para cada unidad con relación a aquellas que están en la

frontera eficiente. Además permite identificar y cuantificar las ineficiencias con relación a los recursos de entrada y los productos de salidas, dando así pautas para el mejoramiento de las distintas unidades analizadas. Esta metodología, basada en Programación Lineal [16].

4. CBR (Case-Based Reasoning). El RBC significa razonar sobre la base de experiencias o "casos" previos. Constituyendo una alternativa entre otras metodologías para construir sistemas basados en el conocimiento. El mismo denota un método donde la solución de un nuevo problema se realiza a partir de las soluciones conocidas para un conjunto de problemas previamente resueltos o no, del dominio de aplicación. Este tipo de técnica, se distingue por utilizar directamente la información almacenada en la memoria del sistema a partir de casos ya resueltos. Para ello utiliza una memoria permanente o base de conocimientos en la cual se almacena de forma explícita la información sobre el dominio de aplicación. Los dominios abordados mediante técnicas RBC implican generalmente tareas de análisis, clasificación, interpretación, diagnóstico, desafíos, planificación o asesoría [17].

Resultados y Objetivos

Este proyecto tiene por objetivo investigar sobre métodos y técnicas en pos de desarrollar un proceso formal que permita asistir a la evaluación y selección de ideas en la plataforma Ideas en tiempo Real, propiciando el uso y la mejora continua de la misma.

Tratándose de un proyecto que recién esta comenzando, los objetivos específicos son los siguientes:

- Investigar acerca los métodos, técnicas requeridas en la evaluación de ideas

- Determinar criterios propicios para de selección de ideas
- Diseñar un proceso formal de evaluación y selección de ideas
- Propiciar que la mayoría de las actividades del proceso de evaluación y selección de ideas se lleven a cabo en un entorno web de tiempo real.
- Determinar los mecanismos necesarios para la gestión de la seguridad del proceso de evaluación y selección de ideas.
- Desarrollar acciones que promuevan el uso y la mejora continua de la plataforma IdeTR.

Formación de Recursos Humanos

El equipo de trabajo que lleva adelante este proyecto se compone de

- 6 docentes investigadores,
- 2 tesistas de grado en período de finalización.
- 1 tesistas de grado en período iniciación.
- 2 tesistas de posgrado (maestría) iniciando sus trabajos.

Referencias

- [1] V. Boly, "Ingénierie de l'innovation. ISBN 97K-2-7462-1798-0," *Lavosier. Paris*, 2008
- [2] D. Díaz, S. Oviedo, A. Otazú, and F. Ibañez, "Desarrollando habilidades para innovar: una experiencia con estudiantes de ingeniería, informática y diseño industrial," *CADI*, 2014.
- [3] D. Díaz, S. Oviedo, C. Muñoz, A. Gallardo, and A. Otazú, "Ideas en tiempo real, un asistente para la generación de ideas en tiempo real," *JATIC 2017-III Jornadas Argentinas de Tecnología Innovación y Creatividad 2017*.
- [4] R. J. M. Tejedor, "WebRTC (Web Real-Time Communications).," *Bit(197)*, 18., 2014.
- [5] Meteor, "What-is-meteor.," 2017.
- [6] S. Rust, J. Schelling, and D. Schipper, "Building real-time web applications with Meteor.."
- [7] G. Salvaneschi, A. Margara, and G. Tamburrelli, "Reactive Programming: A Walkthrough.," *IEEE/ACM 37th IEEE*

- International Conference on Software Engineering.*, 2015.
- [8] Atmosphere.js, "Explore Meteor packages. ," 2017.
- [9] NPM, " Package manager for JavaScript.," 2017.
- [10] E. Keith, " 27 Tools for Online Brainstorming and Decision Making in Meetings.," 2015.
- [11] A. Gabriel, D. Monticolo, and M. Camargo, "Creativity support systems: A systematic mapping study.," *Elsevier, Thinking Skills and Creativity.* , vol. 21, 2016.
- [12] L. V. Gallego, "Tesis:Revision de los metodos, modelos y herramientas existentes para la selection de proveedores. ," *Department of Management and Engineering. Linkopings Universitet.* , 2010.
- [13] M. J. Bach, E. Giupponi, F. S. Sobrero, and M. V. Yori, "Utilización de métodos multicriterio adaptados.," *7mo. Congreso de administración pública, Argentina.* , 2014.
- [14] T. L. Saaty, "Decision making with the analytic hierarchy process. ," *Int. J. Services Sciences*, 2008.
- [15] R. A. Ercole, C. L. Alberto, and C. E. Carignano, "TOPSIS en medición multicriterio de eficiencia.," *XXX CONGRESO ARGENTINO DE PROFESORES UNIVERSITARIOS DE COSTOS, Santa Fe, Argentina.* , 2007.
- [16] Cooper, Seiford, and Tone, "Data Envelopment Análisis," *Edit. Kluwer Academic Publishers.*, 2000.
- [17] D. Cordero Morales, Y. Ruiz Constanten, and Y. Torres Rubio, "Sistema de Razonamiento Basado en Casos para la identificación de riesgos de software. ," *Revista Cubana de Ciencias Informáticas*, 2013.

PLATAFORMA DE GOOGLE: ANDROID MOBILE COMO SOPORTE DE ANDROID THINGS

Mg. Ing. Norma Beatriz Perez⁽¹⁾, Miguel Alfredo Bustos^{(1)(*)}, Dr. Mario M. Berón⁽¹⁾ & PhD. Pedro Rangel Henriques⁽²⁾

⁽¹⁾Departamento de Informática - Facultad de Ciencias Física Matemáticas y Naturales (FCFMyN) - Universidad Nacional de San Luis

⁽²⁾Universidade do Minho - Braga, Portugal

{nbperez, mabustos, mberon}@unsl.edu.ar⁽¹⁾, pedrorangelhenriques@gmail.com⁽²⁾

RESUMEN

En la actualidad, las grandes empresas de desarrollo de software como Google, Apple, Microsoft, Samsung, LG, etc. están abarcando mayor territorio en el mercado global. Es por esto, que Google incorpora al mercado la plataforma de *Android Things* (Android Cosas) que promete un amplio abanico de oportunidades que posibilita a los desarrolladores construir e interconectar gran diversidad de objetos versátiles, seguros, de calidad, de relativo bajo costo, etc. que permiten satisfacer las exigencias de los consumidores.

En este artículo, se describe una línea de investigación que comprende el estudio de la actual y revolucionaria plataforma de *Internet of Things* (Internet de las Cosas, *IoT* siglas en inglés) de Google denominada *Android Things* que está enfocada al desarrollo e implementación de objetos integrados e interconectados. Dicho estudio se fundamenta en la utilización de *Android Mobile* como soporte para construir aplicaciones de *Android Things*. Es por esto, que se lleva adelante un análisis del sistema operativo, arquitectura y seguridad de las actualizaciones de *Android Things*.

Palabras Claves: Android, *Android Things*, *Internet of Things*, Aplicaciones *Mobile*.

CONTEXTO

La presente línea de investigación se enmarca en el Proyecto (PO/16/93) de “*Fortalecimiento de la Seguridad de los Sistemas de Software mediante el uso de*

Métodos, Técnicas y Herramientas de Ingeniería Reversa”. Realizado en conjunto con la Universidade do Minho Braga, Portugal. Recientemente aprobado por el Ministerio de Ciencia Tecnología e Innovación Productiva (Mincyt). Y por el Proyecto (P031516.) de Investigación: “*Ingeniería de Software: Conceptos, Prácticas y Herramientas para el Desarrollo de Software con Calidad*”. De la Facultad de Ciencias Físico Matemáticas y Naturales, Universidad Nacional de San Luis. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el Programa de Incentivos.

1. INTRODUCCIÓN

Las tecnologías móviles han revolucionado la sociedad, en todos los ámbitos como el académico, industrial, militar, social, entre otros, de manera significativa como lo ha hecho Internet. El lanzamiento de Android [1, 2] como plataforma para el desarrollo de aplicaciones móviles [3] ha causado gran impacto a nivel mundial. Esto a logrando gran aceptación por parte de sus usuarios como así también por la industria en general. En la actualidad, la plataforma Android [4] se ha convertido en una alternativa dominante frente a otras plataformas. Dicha plataforma, es de código abierto y está basada en el sistema operativo Linux lo que permite la utilización de la misma sin costo adicional, es simple y adaptable a cualquier tipo de hardware como

por ejemplo, Smartphone, Tablet, Smartwatch, SmartTV, automóviles, electrodomésticos y una amplia variedad de productos empotrados que utilizan este sistema operativo para llevar a cabo sus tareas y programaciones. Además, la plataforma Android provee portabilidad segura dado que las aplicaciones son desarrolladas en lenguaje JAVA [5] lo que permite que las aplicaciones sean ejecutadas en cualquier dispositivo a través de la máquina virtual *Dalvik* [6]. Por otro lado, la arquitectura de Android [7] se basa en componentes “*inspirados*” en Internet, como por ejemplo, la interfaz de usuario se realiza en código XML lo que permite que una misma aplicación se ejecute en diferentes pantallas con distintas dimensiones como por ejemplo un dispositivo pequeño como un reloj o en un televisor. Android incorpora servicios de localización basado en GPS [8], en redes, bases de datos con SQL, síntesis de voz, multimedia, etc. La seguridad de Android [9] se provee a través de permisos que son otorgados por parte de sus usuarios.

Android Mobile, ofrece una forma sencilla y novedosa de implementar y desarrollar potentes aplicaciones para diferentes tipos de dispositivos.

Por otro lado, **Android Things** [10] es una plataforma que extiende de *Android Mobile* con el objetivo de proporcionar a los desarrolladores la posibilidad de construir objetos integrados e interconectados con características destacables como la alta calidad, mayor seguridad, productos a escala, etc. La plataforma de *Android Things* se compone de tres pilares fundamentales, con respecto a *Android Mobile*, que se describen a continuación.

- Optimización del Sistema Operativo [11]: Una variante para ser utilizada en *IoT* [12].
- Arquitectura: Se extiende el marco central con APIs adicionales proporcionadas por la biblioteca soporte de *Things*.
- Actualizaciones seguras: Administradas directamente por Google.
- Hardware potente: Accesible y de fácil integración.

La línea de investigación aborda el estudio y análisis de los pilares principales (ítems) de *Android Things* mencionados previamente. Así como las componentes que extienden de *Android Mobile* que llevan a la transformación de *Android Things*.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Android Things facilita el desarrollo de dispositivos integrados conectados a través de la utilización de las mismas herramientas de desarrollo de *Android Mobile* como *Android Studio* [13, 14], las APIs, los recursos y la enorme comunidad de desarrolladores expertos en Android.

A continuación se describen los principales componentes de *Android Things*.

2.1 SO Android Things

El software que se ejecuta en el dispositivo con *Android Things*, permite construir aplicaciones que utilizan el marco proporcionado por *Android Mobile*, kit de desarrollo de software (SDK) y servicios de Google Play [15]. Esto incluye la misma interfaz de usuario, toolkit, soporte multimedia y APIs de conectividad utilizadas por los desarrolladores de *Mobile*. Estas aplicaciones se integran fácilmente con los servicios más habituales de Google, como *Firebase* [16], *TensorFlow* [17] y *Google Cloud Platform* [18] utilizando la diversidad de bibliotecas de desarrolladores de *Android Mobile*.

El desarrollo de *Android Things* utiliza el mismo lenguaje y herramientas que se utiliza para construir en *Android Mobile*. Sin embargo, se ha ajustado la plataforma para reducir los tiempos de arranques como así también reducir el consumo de memoria incluyendo una variante de servicios de Google Play optimizada para *IoT*. Además, se ha agregado nuevas APIs a fin de integrarse adecuadamente con el hardware personalizado; interfaces periféricas y administración de dispositivos. Por otro lado, *Android Things* no dispone de aplicaciones de

usuario como un navegador o indicador. Esto significa que está diseñado para comenzar directamente con las aplicaciones que se han creado para el dispositivo.

2.2 Arquitectura Android Things

Android *Things* amplía el marco central de Android con APIs adicionales proporcionadas por la biblioteca de soporte de *Things* que le permite integrarse con nuevos tipos de hardware que no se encuentran en los dispositivos móviles tradicionales.

El desarrollo de aplicaciones para dispositivos integrados es diferente del móvil en algunos aspectos importantes que se mencionan a continuación.

- Acceso más flexible a periféricos de hardware y controladores.
- Las aplicaciones del sistema no están presentes para optimizar los requisitos de inicio y almacenamiento.
- Las aplicaciones se inician automáticamente al inicio para sumergir a los usuarios en la experiencia de la aplicación.
- Los dispositivos muestran sólo una aplicación a los usuarios, en lugar de múltiples como los dispositivos móviles.

En la Figura N° 1 se observa la arquitectura de Android *Things*.



Figura N° 1: Arquitectura Android *Things*.

2.3 Actualizaciones Android Things

Google proporciona actualizaciones y parches de seguridad para el sistema operativo central a fin de que el desarrollador se pueda enfocar específicamente en la construcción de la aplicación. Además, esto permite mantener protegidos a los usuarios en todo momento. Por otro lado, las imágenes

del sistema están firmadas por Google y verificadas por integridad en el dispositivo lo que evita una actualización corrupta o alterada. En caso de producirse un error en una actualización, el sistema iniciará en un estado conocido previo donde se encuentre estable.

Las actualizaciones se envían por Internet desde la consola Android *Things* utilizando la misma infraestructura segura que se usa para actualizar los dispositivos móviles en la actualidad. Además, previene las actualizaciones automáticamente cuando los parches de seguridad están disponibles para la plataforma.

Las aplicaciones en el dispositivo se administran exclusivamente a través de la consola *Things* e incluye cada actualización, por lo que Android *Things* no incorpora Google Play Store ya que las aplicaciones instaladas por el usuario no son soportadas.

Android *Things Console* [19] (Ver Figura N° 2) proporciona herramientas para instalar y actualizar la imagen del sistema en dispositivos de hardware compatibles. Esto permite enviar actualizaciones a los usuarios así como probar las implementaciones en su propio hardware. La utilización de la consola proporciona:

- Descargar e instalar la última imagen del sistema Android *Things*.
- Crear imágenes de fábrica que contengan aplicaciones OEM junto con la imagen del sistema.
- Lanzar las actualizaciones por aire (OTA), incluidas las aplicaciones OEM y la imagen del sistema, a los dispositivos.

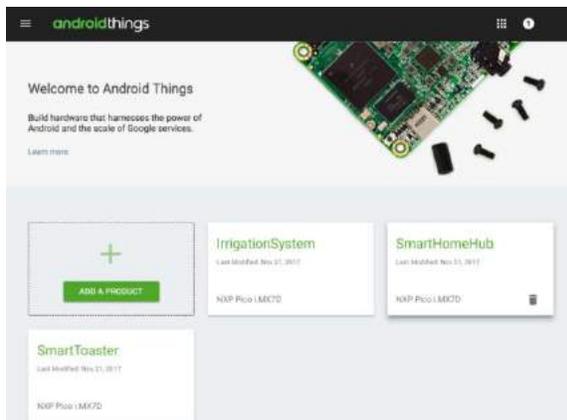


Figura N° 2: Consola de Android Things.

2.4 Hardware Android Things

El hardware Android Things está pre-certificado por Google con un paquete de soporte proporcionado por proveedores. Dichos proveedores brindan soporte a largo plazo para cada lanzamiento estable en toda línea de hardware. El hardware también está pre-certificado con agencias reguladoras como la FCC [20], simplificando las pruebas de certificación requeridas como la de llevar su producto al mercado y poder así ahorrar tiempo y dinero por parte de los desarrolladores.

Para el desarrollo de una aplicación con Android Things es necesario disponer una placa soporte con periféricos como por ejemplo: *NXP Pico i.MX7D*, *NXP Argon i.MX6UL*, *Raspberry Pi 3 Model B* [21], etc.

3. RESULTADOS OBTENIDOS/ESPERADOS

El trabajo de investigación permitió obtener diferentes resultados que serán utilizados para poder transformar aplicaciones de Android Mobile a aplicaciones de Android Things con fines similares. A continuación se describe de manera sucinta algunos análisis comprendidos en la línea de investigación.

- Se determinó que Android Things es una extensión de Android Mobile lo que permite transformar aplicaciones realizadas y utilizadas actualmente a aplicaciones Things.

- Android Things permite arrancar una aplicación a la vez en cada dispositivo lo que permite unificar velocidades de procesamiento como así también minimizar la utilización de memoria.
- Android Things extiende la arquitectura de Android Mobile a fin de incorporar APIs proporcionadas por la biblioteca de IoT.
- Android Things utiliza una consola para realizar actualizaciones seguras en los dispositivos. Dichas actualizaciones son administradas por la central de Google.
- Para el desarrollo de Android Things se utiliza la misma plataforma de desarrollo que se usa tradicionalmente para el desarrollo de aplicaciones Mobile. Este entorno de desarrollo integrado se denomina *Android Studio*. Android Studio incorpora las extensiones de Things necesarias para el desarrollo de dispositivos integrados y conectados.

Los investigadores de esta línea continuarán con estudios en este campo con el objetivo de perfeccionarse en el área y realizar transformaciones de aplicaciones tradicionales a través del uso de la plataforma Android Things.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de profesionales de la UNSL que forman parte de la línea de investigación de este trabajo llevan adelante diferentes trabajos finales integradores de *Ingeniería en Informática*, *Ingeniería en Computación*, *Licenciatura en Ciencias de la Computación*, y en un futuro próximo trabajos finales de especialización, tesis de maestría y doctorado. En particular, las investigaciones desarrolladas en este trabajo forman parte del lineamiento inicial como trabajo final de uno de los autores para optar al grado de Ingeniero en Informática en la UNSL.

5. BIBLIOGRAFÍA

- [1] Burnette, E. (2015). *Hello, Android: introducing Google's mobile development platform*. Pragmatic Bookshelf.
- [2] Tomás Gironés, J. (2016). *El gran libro de Android*. 4ª. Edición. Alfaomega.
- [3] Bustos, M. A., Perez, N. B., & Berón, M. (2015, May). *Plataformas para el desarrollo de aplicaciones móviles*. In *XVII Workshop de Investigadores en Ciencias de la Computación* (Salta, 2015).
- [4] Sitio oficial de Android: <https://www.android.com/>
- [5] Kurniawan, B. (2015). *Java for Android*. Brainy Software Inc.
- [6] Zambrano, G. R. (2016). Análisis de Comparación de Android y GNU/Linux/Comparison Analysis Android and GNU/Linux. *International Journal of Innovation and Applied Studies*, 18(4), 1039.
- [7] Londoño, S., Urcuqui, C. C., Cadavid, A. N., Amaya, M. F., & Gómez, J. (2015). SafeCandy: System for security, analysis and validation in Android. *Sistemas & Telemática*, 13(35), 89102.
- [8] Sultana, S., Enayet, A., & Mouri, I. J. (2015). A Smart, Location Based Time and Attendance Tracking System Using Android Application. *International Journal of Computer Science, Engineering and Information Technology (IJCEIT)*, 5(1), 15.
- [9] Elenkov, N. (2014). *Android security internals: An indepth guide to Android's security architecture*. No Starch Press.
- [10] Sitio oficial Android Things: <https://developer.android.com/things/index.html>
- [11] Hahm, O., Baccelli, E., Petersen, H., & Tsiftes, N. (2016). Operating systems for lowend devices in the internet of things: a survey. *IEEE Internet of Things Journal*, 3(5), 720734.
- [12] Wortmann, F., & Flüchter, K. (2015). Internet of things. *Business & Information Systems Engineering*, 57(3), 221224.
- [13] DiMarzio, J. F. (2015). Setting Up Android Studio. In *Android Studio Game Development* (pp. 18). Apress, Berkeley, CA.
- [14] Sitio oficial Android Studio: <https://developer.android.com/studio/index.html>.
- [15] Viennot, N., Garcia, E., & Nieh, J. (2014, June). A measurement study of google play. In *ACM SIGMETRICS Performance Evaluation Review* (Vol. 42, No. 1, pp. 221233). ACM.
- [16] Sitio oficial Firebase: <https://firebase.google.com/>
- [17] Sitio oficial TensorFlow: <https://www.tensorflow.org/>
- [18] Sitio oficial Google Cloud Platform: <https://cloud.google.com/>
- [19] Sitio oficial de desarrollo Android Things: <https://developer.android.com/things/console/index.html>
- [20] Sitio oficial de Federal Communications Commission: <https://www.fcc.gov/consumers/guides/acercadelafcc>
- [21] Pi, R. (2015). Raspberry Pi Model B.

Rampas Digitales Innovativas para Personas con Discapacidad

Ivana Harari, Laura Fava, Javier Díaz, Paula Altoaguirre, Rodrigo Torales
LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.

Facultad de Informática. Universidad Nacional de La Plata

Calle 50 esq. 120, 2do Piso. Tel: +54 221 4223528

{iharari, lfava, jdiaz}@info.unlp.edu.ar, {rodrigotorales,paulialtolaguirre}@gmail.com

Resumen

Según la OMS, Organización Mundial de la Salud, el 15 por ciento de la población mundial está afectado por alguna discapacidad física, psíquica o sensorial que dificulta su desarrollo personal y su integración social, educativa o comunicacional. Las tecnologías de apoyo o rampas digitales hacen referencia a dispositivos, equipos, instrumentos, recursos tecnológicos o software que permiten incrementar, optimizar o suplir ciertas capacidades funcionales de las personas con discapacidad, que se encuentran limitadas o ausentes.

En este artículo se describen las actividades que se están desarrollando con y para personas ciegas, con el objetivo de diseñar y desarrollar rampas digitales innovativas, a nivel de software y a nivel de hardware, que asistan a las personas ciegas en sus actividades cotidianas. Específicamente, se propone investigar y desarrollar un kit de anteojos y bastón wearables, basados en sensores, para el tránsito y circulación de personas ciegas.

Palabras claves: sensores, wearables, discapacidad visual, accesibilidad, tecnologías asistivas, Internet of Things.

Contexto

La Facultad de Informática, a través de la Dirección de Accesibilidad y del Laboratorio de Investigación de Nuevas Tecnologías Informáticas LINTI, ha llevado a cabo en los últimos años, líneas de acción concretas que estrechan el vínculo Facultad-Sociedad, atendiendo las demandas de los sectores más vulnerables de la comunidad, como lo son las personas con discapacidad. A través de varios cursos de formación abierto a todo público como el de Accesibilidad Web, el de TICs para personas con discapacidad; de jornadas y conferencias como la de una Facultad Inclusiva, la de Experiencias y Casos de Aplicación de desarrollos accesibles, las hackatones de 24 hs. de desarrollo por la discapacidad; como también, a través de la dirección de proyectos sobre inclusión social y accesibilidad hace que la Facultad de Informática se constituya como una entidad de referencia respecto a cuestiones de tecnología y discapacidad.

Asimismo en el LINTI se viene trabajando en la construcción de dispositivos basados en sensores, los cuales prometen transformar muchos aspectos de la forma en que vivimos. En este sentido se propone la utilización de estas nuevas tecnologías para la elaboración de wearables que ayuden a las personas con ceguera a transitar de

manera más independiente y segura.

El proyecto que se describe en este artículo realiza un abordaje de esta problemática mediante actividades de investigación, innovación, desarrollo y de búsqueda permanente de soluciones tecnológicas que mejoren la calidad de vida de las personas que atraviesan alguna discapacidad.

El mismo se desarrolla en el Laboratorio de Investigación en Nuevas Tecnologías Informáticas, LINTI de la Facultad de Informática de la UNLP y está enmarcado en los proyectos 11-F014 *Innovación en TICs para el desarrollo de aplicaciones en educación, inclusión, gobierno y salud* finalizado en 2015 y en el nuevo proyecto *Internet del futuro: Ciudades digitales inclusivas, innovadoras y sustentables, IoT, ciberseguridad y espacios de aprendizaje del futuro*, ambos acreditados en el marco del Programa de Incentivos, bajo la dirección del Lic. Javier Díaz.

Introducción

Las nuevas tecnologías están permitiendo que las cosas se conecten en cualquier momento, en cualquier lugar, con cualquier otra cosa, usando idealmente cualquier red y cualquier servicio, dando lugar a lo que se conoce como Internet de las Cosas (IoT).

IoT permite crear entornos habilitantes ofreciendo asistencia a personas con discapacidades en la construcción de acceso, transporte, información y comunicación. IoT puede ser aplicado a múltiples escenarios para facilitar que personas con impedimentos puedan llevar a cabo sus ocupaciones diarias (Domingo, M., 2012). Esto aumenta su autonomía y confianza en sí mismo, ser independiente en las actividades diarias de uno sin

requerir la asistencia de una persona vidente es la más alta prioridad para personas con discapacidad visual (Lanigan et al., 2007).

Este artículo presenta una línea de investigación cuyo objetivo es desarrollar de rampas digitales innovativas basadas en sensores para que personas ciegas optimicen su movilidad, tránsito y comunicación social. Se propone, más específicamente, desarrollar un kit conformado por un bastón y anteojos optimizados con sensores y una aplicación móvil multiplataforma para asistir a las personas ciegas para un tránsito más autónomo por la ciudad.

Líneas de Investigación, Desarrollo e Innovación

Las líneas de investigación, desarrollo e innovación que se llevan a cabo en este proyecto son:

- Análisis de las diferentes tecnologías y desarrollo de soluciones para personas ciegas.
- Construcción de las rampas digitales basadas en sensores que permitan mejorar la vida de personas con esta discapacidad.
- Estudio de la interacción con personas ciegas teniendo en cuenta la capa de accesibilidad de los sistemas operativos de los dispositivos móviles.
- Análisis de la interfaz de programación brindada por los ómnibus públicos e integración con la solución propuesta.

Resultados y Objetivos

Como se mencionó anteriormente, en el LINTI, se han desarrollado muchas herramientas que permiten mejorar la

calidad de personas con discapacidad. Muchas de estas herramientas han sido analizadas y, constituyeron el insumo y motivación para investigar sobre nuevas rampas digitales como el desarrollo del kit de asistencia para la circulación para personas ciegas y con disminución visual.

Entre las herramientas más importantes, se puede mencionar a: TalkLouder, Liberium, Tracks TransitaWeb, Navegación Web asistida por comandos de voz, calzado háptico.

Comenzaremos con TalkLauder, una aplicación móvil que le permite a las personas sordas o hipoacúsicos comunicarse con personas oyentes, se trata de una aplicación innovadora, la cual está orientada a un público específico, particularmente los usuarios con algún grado de discapacidad auditiva que mediante la selección de frases precargadas y categorizadas les permite fácilmente comunicarse.

Otra iniciativa vinculada con accesibilidad es Liberium, una aplicación para gestionar lugares de interés público con información, en especial, su nivel de accesibilidad. Está dirigida en un principio a la ciudad de La Plata y permite a las personas con movilidad reducida, planificar sus salidas para no encontrarse con barreras al momento de arribar a un lugar (Ibáñez, L., 2016).

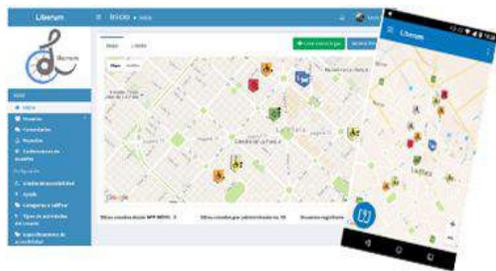


Fig 1: Liberium

Esta aplicación es una convocatoria abierta a la comunidad de La Plata para colaborar con la construcción de una ciudad más accesible e inclusiva.

Tracks es otra aplicación móvil, para el seguimiento conjunto de pacientes con discapacidad, donde los sujetos intervinientes en su tratamiento como ser médico, terapeuta, docente, psicólogo, tutores, padres y otros profesionales, puedan registrar situaciones, brindar observaciones, informar datos relevantes al seguimiento y evolución del tratamiento, y configurar su propagación a quienes corresponda (Trejo, M., Vilas, L., 2016). Provee estrategias efectivas de comunicación conjunta, de trabajo colaborativo, como también el acceso en línea a la información integral y actualizada, sobre la persona con discapacidad a la que se encuentra en tratamiento.

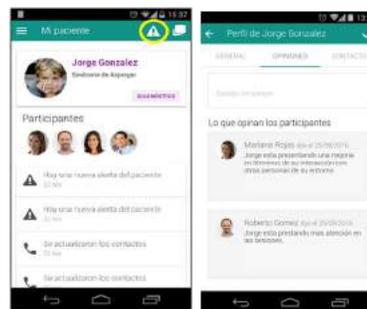


Fig 2: Tracs

TransitaWeb es una aplicación Web que convierte el sitio que la persona con discapacidad quiera acceder, a un sitio accesible y adaptado de acuerdo a las necesidades del usuario. Realiza un proceso de accesibilización aplicando en forma automática las normas WCAG 2.0 (W3C, 2008). Presenta una página accesible donde el usuario con discapacidad indica el sitio a visitar y además puede configurar las adaptaciones que requiera, como ser el tamaño de letra

que necesita, los colores que le resultan más adecuados, organización estructural de la página que mejor le resulte, como también el tipo de contenidos que quisiera descartar, como ser publicidades, accesos a redes sociales, paneles secundarios, entre otros. En la siguiente figura se muestra la página principal de TransitaWeb:

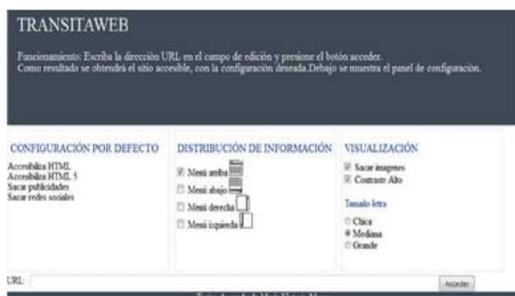


Fig 3: TransitaWeb

Por otro lado, se encuentra la herramienta que permite la navegación de sitios por comandos de voz posibilitando al usuario con discapacidad utilizar el recurso del habla para navegar, seleccionar como acceder a los contenidos de los sitios, sin necesidad de utilizar el teclado ni el mouse. Esta herramienta fue desarrollada como una extensión o complemento del navegador, y permite la ejecución de comandos de voz mediante Web Speech API (Pérez J., 2017).

Otra iniciativa es el calzado háptico, una solución tecnológica integral de bajo costo en forma de calzado háptico y una aplicación móvil para guiar a ciegos a un destino específico en tiempo real que también permite la detección de obstáculos físicos durante el recorrido (Berretti, F., 2014).



Fig 4: Calzado Háptico

Finalmente, un trabajo vinculado con sensores y discapacidad, que está en etapa de finalización, es la integración de softwares aumentativos y alternativos con interfaces cerebro-computadora (Díaz, J., 2016), para potenciar la autonomía personal y la calidad de vida de personas con discapacidad agudas.

La cantidad de trabajos descriptos desarrollados en el LINTI, la experiencia y conocimiento adquirido, da un marco teórico y experimental para proponer este kit basado en sensores para personas ciegas. Por otra parte, además de esto se han profundizado los relevamientos de las problemáticas de las personas con ceguera, donde se han analizado las demandas y necesidades respecto a movilidad, comunicación y manejo de recursos, que se ven limitados por la escasa o nula visión. También se han efectuado entrevistas, y realizado encuestas como otros recursos de indagación, de observación y de registro, para la confección conjunta de los requisitos funcionales de las posibles soluciones tecnológicas a desarrollar.

Específicamente los objetivos de este proyecto son:

- Analizar los tipos de sensores apropiados para la creación de un bastón que detecte objetos en área baja y superior.

- Determinar qué tipo de sensor es el más apropiado para la creación de anteojos que detecten la proximidad de objetos.
- Diseñar y desarrollar una aplicación móvil que configure y administre de manera simple el kit creado.
- Analizar la información brindada por los GPS de los ómnibus públicos para informar mediante voz la ubicación de un ómnibus requerido por la persona ciega, desde el momento en que hace la consulta hasta el momento en que el ómnibus llega a la parada.
- Probar los prototipos de hardware y de software desarrollados con personas ciegas y disminuidos visuales en distintos contextos de interacción.
- Analizar los resultados obtenidos y valoración del impacto de su uso teniendo en cuenta conectividad, eficiencia, nivel de aceptación, grado de satisfacción.

Formación de Recursos Humanos

El equipo de trabajo de la línea de I+D+i presentada en este artículo está formado por docentes investigadores categorizados del LINTI y estudiantes de la Facultad de Informática.

Se encuentran dos tesis en desarrollo vinculadas con esta temática vinculada con los desarrollos propuestos.

A través de la generación permanente de conocimiento por medio de esta línea de investigación y desarrollo de aplicaciones para accesibilidad, el LINTI promueve el uso innovador de las tecnologías informáticas en la región.

Referencias

Berretti, F. *Calzado Háptico: Navegabilidad Asistida para Personas con Disminución Visual*. Tesis de Grado accesible en <http://sedici.unlp.edu.ar/handle/10915/63352>. Diciembre 2014.

Domingo, M. (2012) *An overview of the Internet of Things for people with disabilities*, Journal of Network and Computer Applications, Electrical Engineering, UPC-Barcelona Tech University, Barcelona, España.

Ibáñez, L. (2016) *Liberum: aplicación móvil para socializar la accesibilidad en la ciudad de La Plata*. Tesis de Grado disponible en <http://sedici.unlp.edu.ar/handle/10915/58717>

Lanigan P., Paulos A., Williams A., Rossi D, Narasimhan P. *Trinetra: assistive technologies for grocery shopping for the blind*. In: Proceedings of the IEEEBAIS symposium on research in assistive technologies. Dayton, USA; April 2007.

Pérez, J., Díaz, J., Harari, I. *Ejecución de comandos de voz mediante Web Speech API*. CACIC 2017, XXIII Congreso Argentino de Ciencias de la Computación. ISBN: 978-950-34-1539-9. La Plata, Octubre 2017.

Trejo, M., Vilas, L., Harari, I., Amadeo, P. *TraCS: Trabajo, Colaboración y Seguimiento: herramienta de seguimiento colaborativa de tratamientos de pacientes con discapacidad-* (2016) Tesis de Grado accesible en <http://sedici.unlp.edu.ar/handle/10915/59013>

W3C- W3 Consortium. *Web Content Accessibility Guidelines WCAG 2.0*. Disponible en <https://www.w3.org/TR/WCAG20/>

Agentes Inteligentes y Web Semántica: Verbalización en una herramienta Web de modelado ontológico

Matías Garrido¹Germán Braun^{1,2,3}Sandra Roger¹

email: {roger,german.braun}@fi.uncoma.edu.ar

¹Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

²Laboratorio de I&D en Ingeniería de Software y Sistemas de Información
Departamento de Ciencias e Ingeniería de la Computación
UNIVERSIDAD NACIONAL DEL SUR

³Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

RESUMEN

El proyecto de investigación Agentes Inteligentes y Web Semántica, financiado por la Universidad Nacional del Comahue, tiene como objetivo general la generación de conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la Web, es decir, lo que se ha llamado la Web Semántica.

El objetivo general del trabajo de investigación es la extensión de una herramienta de modelado ontológico, denominada crowd, mediante la verbalización de un subconjunto del lenguaje de modelado conceptual UML. Esta integración permitirá generar especificaciones en Lenguaje Natural a partir de un diagrama de clases.

Esta línea de investigación se desarrolla en forma colaborativa entre docentes-investigadores de la Universidad Nacional del Comahue y de la Universidad Nacional del Sur, en el marco de proyectos de investigación financiados por las universidades antes mencionadas.

Palabras Clave: Verbalización, Generalización de Lenguaje Natural, UML, Ontologías.

CONTEXTO

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el contexto del proyecto de investigación Agentes Inteligentes y Web Semántica (04/F014), por la Universidad Nacional del Sur a través del proyecto de investigación Integración de Información y Servicios en la Web (24/N027) y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral. Los proyectos de investigación tienen una duración de cuatro años y la beca doctoral con duración de 5 años, finalizando en abril de 2019.

1. INTRODUCCIÓN

Según Christel et al. [1], la etapa inicial de un proceso de desarrollo de software es la más problemática. La complejidad radica en la correcta elicitación de los requerimientos del sistema, en su entendimiento y en la transformación de éstos en un modelo computacional.

Cualquier error cometido en estas instancias se propagará hacia las fases posteriores y su corrección incrementará los costos de desarrollo, especialmente si se identifican

durante la fase de implementación o posterior a la entrega del producto.

A medida que el proceso de desarrollo transcurre, dichos requerimientos deben ser trasladados hacia una especificación formal. Curland et al. [2] afirma que es importante que un esquema conceptual sea validado por el usuario con el objetivo de asegurar que ésta modela con precisión los aspectos relevantes del dominio. Una manera efectiva de facilitar esta validación es la traducción de tal esquema (o cualquier especificación formal) en un lenguaje claro que sea fácilmente comprensible por el experto en el dominio, quien, además, puede no poseer conocimientos en modelado de sistemas.

En esta dirección surge el concepto de verbalización, el cual consiste en traducir la semántica de una teoría lógica, es decir, las relaciones entre entidades y sus restricciones, en sentencias en lenguaje natural [3][4]. En consecuencia, esta transformación proporciona una base para la mutua comprensión entre diferentes actores acerca de la semántica que una lógica capta (ya sea un modelo conceptual, reglas de negocio o una ontología), integrando clientes/usuarios en los procesos de chequeo de consistencia cuando los cambios son realizados en el diseño o en la implementación.

La verbalización es un proceso que tiene sus bases y está comprendido dentro del campo de la Generación de Lenguaje Natural (GLN). Reiter y Dale [5] lo definen como al ámbito de la inteligencia artificial y la lingüística computacional que se ocupa de la construcción de sistemas informáticos que pueden producir textos comprensibles en español, inglés u otros lenguajes naturales de alguna representación no lingüística de información.

McDonald [6], definió a la GLN como el proceso de construir un texto en lenguaje natural para cumplir objetivos comunicativos específicos. Estos sistemas tienen el desafío no sólo de decidir qué información comunicar sino cómo transformarla, de modo que se alcance la meta de comunicación deseada.

Sin embargo, poco se ha investigado en validar los modelos de software de una manera en que los usuarios puedan comprenderlos.

Existen varias herramientas de GLN a tener en cuenta, tales como ModEx (Model Explainer [7] y GeNLangUML (*Generating Natural Language from UML*) [8]. Ambas permiten la creación y edición gráfica de modelos de software orientados a objetos (OO) [9] y diagramas de clases UML [10].

respectivamente, para generar especificaciones siendo el idioma de verbalización el inglés.

Mientras que ModEx supone que los nombres de las clases como de las relaciones de un diagrama OO son

correctos y cumplen con ciertas recomendaciones, GeNLangUML tiene la ventaja de utilizar

WordNet [11], una herramienta lingüística para validar la correctitud semántica de las sentencias generadas.

Esto implica que funcione bajo una convención de nomenclatura

más amplia para los componentes de un diagrama de clases.

Con respecto a la verbalización multilingüe, y en relación al lenguaje ORM [12], DogmaModeller [3] es una aplicación que genera automáticamente verbalizaciones en lenguaje pseudo-natural. Un aspecto a destacar de esta herramienta es su extensibilidad, es decir, la capacidad de incorporar nuevos templates de verbalización para otros idiomas.

Otra herramienta a tener en cuenta, que soporta modelos ORM de segunda generación (ORM2) [13], es NORMA (Neumont ORM Architect) [2].

Por otro lado, KeY [14] es un sistema que ayuda a comprender las restricciones OCL [15], traduciéndolas automáticamente en lenguaje natural.

La investigación desarrollada en [16] presenta un algoritmo que proporciona parafraseos en lenguaje natural para Ontologías OWL [17] en la Web Semántica (WS) [18]. Bajo la misma dirección, se encuentra NaturalOWL [19]. Generalmente los sistemas de GLN agregan el máximo de frases posibles para mejorar la legibilidad, pero NaturalOWL posee la ventaja de configurar el número máximo de frases que se desean concatenar.

2. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

El proyecto de investigación Agentes Inteligentes y Web Semántica tiene como objetivo general generar conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la web, es decir lo que se ha llamado la Web Semántica.

Por otro lado, en el proyecto de investigación Integración de Información y Servicios en la Web se propone investigar y desarrollar metodologías y herramientas que favorezcan la interoperabilidad semántica de información y de servicios en la Web, fundamentados en los últimos avances en el área de lenguajes de representación del conocimiento, ontologías y modelado conceptual.

Ambos proyectos confluyen en la línea de investigación de este trabajo, en la que se explora entre otros, sobre temas afines a la Representación del Conocimiento, las Lógicas Descriptivas, [20], las Ontologías, la Ingeniería de Software basada en Conocimiento y la Ingeniería de Conocimiento.

En base a las investigaciones existentes hasta la fecha, se propone modificar la arquitectura de una herramienta de modelado ontológico llamada *crowd* [21], mediante la incorporación de una nueva funcionalidad que soporte verbalización multilingüe del metamodelo UML, orientado hacia un subconjunto de primitivas de los diagramas de clases. Se pretende la utilización de una representación intermedia de los diagramas de clases en sentencias concretas de lógica de primer orden (FOL) para favorecer las verbalizaciones en diversos idiomas.

Como principal novedad con respecto a las aplicaciones vigentes, se puede destacar que los usuarios podrán visualizar y editar de forma *on-line* diagramas de clases UML y posteriormente, realizar su verbalización a lenguaje natural. Esto se debe a que *crowd* es soportada por una arquitectura cliente-

servidor, lo que conlleva a que pueda ser utilizada bajo cualquier plataforma.

Otro aspecto relevante, es que la traducción a lenguaje natural está destinada a ser multilingüe, a diferencia de la mayoría de las herramientas mencionadas anteriormente, donde el lenguaje natural de salida es preferentemente el inglés.

3. RESULTADOS OBTENIDOS/ESPERADOS

Este trabajo forma parte de la propuesta presentada inicialmente en [22]. Se trabajó en el análisis y diseño de la extensión de la arquitectura de la herramienta *crowd*. Se diseñó el módulo de verbalización de dicha arquitectura y se está trabajando en la fase de implementación del mismo.

Se trabajó en la selección del subconjunto de primitivas en UML destinadas a la verbalización en una primera etapa del proceso. Además, se estableció la correspondencia directa entre estas primitivas y las sentencias de FOL.

Actualmente se está trabajando en la fase de verbalización de esas sentencias FOL. Posteriormente se prevee la extensión a otros idiomas y la correspondiente validación.

4. FORMACIÓN DE RECURSOS HUMANOS

Durante la realización de este sistema se espera lograr, como mínimo, la culminación de 2 tesis de grado dirigidas y/o codirigidas por los integrantes del proyecto. Uno de los autores de este trabajo está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur (beca interna doctoral CONICET).

Finalmente, es constante la búsqueda hacia la consolidación como investigadores de los miembros más recientes del grupo.

5. BIBLIOGRAFÍA

- [1] Michael Christel and Kyo Kang. Issues in requirements elicitation. Technical Report CMU/SEI-92-TR-012, Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, 1992.

- [2] Matthew Curland and Terry A. Halpin. The NORMA software tool for ORM 2. In Information Systems Evolution - CAiSE Forum 2010, Hammamet, Tunisia, June 7-9, 2010, Selected Extended Papers, pages 190–204, 2010.
- [3] Mustafa Jarrar. Towards Methodological Principles for Ontology Engineering. PhD thesis, Vrije Universiteit Brussel, Brussels, 5 2005.
- [4] Mustafa Jarrar, María Keet. Multilingual verbalization of ORM conceptual models and axiomatized ontologies. Vrije Universiteit Brussel. Brussels, 2006.
- [5] Reiter Euhand, Robert Dale. Building Applied Natural Language Generation Systems. Cambridge University Press. NJ, USA, 1989.
- [6] McDonald David, Gabriel Richard. Natural Language Generation Systems. 1988.
- [7] Benoit Lavoie, Owen Rambow, and Ehud Reiter. The modelexplainer, 1996.
- [8] Farid Meziane, Nikos Athanasakis, and Sophia Ananiadou. Generating natural language specifications from UML class diagrams. *Requir. Eng.*, 13(1):1–18, 2008.
- [9] Won Kim and Frederick Lochovsky. Object-oriented concepts, databases, and applications. ACM Press/Addison-Wesley Publishing Co 1989.
- [10] UML, OMG. Unified Modeling LanguageTM. 2015.
- [11] George Miller. Wordnet: a lexical database for English. ACM.1995.
- [12] Terry Halpin and Tony Morgan. Information Modeling and Relational Databases. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2 edition, 2008.
- [13] Robert Meersman and Zahir Tari. On the Move to Meaningful Internet Systems 2005: {OTM} 2005 Workshops, {OTM} Confederated International Workshops and Posters. 2005
- [14] Ahrendt, Wolfgang and Baar, Thomas and Beckert. The KeY tool. 2005.
- [15] Warmer, Jos and Kleppe, Anneke. The Object Constraint Language: Precise Modeling with UML. 1999.
- [16] Daniel Hewlett and Aditya Kalyanpur and Vladimir Kolovski and Christian Halaschek-Wiener. Effective NL Paraphrasing of Ontologies on the Semantic Web. 2005.
- [17] McGuinness, Deborah L. and van Harmelen, Frank. OWL Web Ontology Language Overview. W3C. 2004.
- [18] The semantic web. Berners-Lee, Tim and Hendler, James and Lassila, Ora and others. New York, NY, USA. 2001
- [19] Ion Androutsopoulos and Gerasimos Lampouras and Dimitrios Galanis. Generating Natural Language Descriptions from OWL Ontologies: the NaturalOWL System. 2014.
- [20] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. The Description Logic Handbook: theory, implementation, and applications. 2003.
- [21] Christian Gimenez, Germán Braun, Laura Cecchi, and Laura Fillottrani. crowd: A Tool for Conceptual Modelling assisted by Automated Reasoning - Preliminary Report. In Proc. of the 2nd Simposio Argentino de Ontologías y sus Aplicaciones (SAOA) colocated at

Jornadas Argentinas de Informática (JAIIO) - to appear, 2016.

- [22]Matías Garrido, Germán Braun, Sandra Roger. Agentes Inteligentes y Web Semántica: Hacia la Verbalización de un Subconjunto de UML en una Herramienta Gráfica Web. XIX Workshop de Investigadores en Ciencias de la Computación (WICC). 2017.

Procesamiento Distribuido y Paralelo

Orquestación de servicios para la Continuidad Edge al Cloud

Nelson Rodríguez¹, María Murazzo¹, Susana Chávez¹, Adriana Martín¹, Diego Medel¹, Jorge Mercado², Miguel Guevara¹, Fabiana Píccoli³, Miguel Méndez Garabetti⁴, Javier Rosentein⁴, Mónica Gimenez⁵, Javier Sillero⁶

¹ Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.

² Departamento de Matemática - Facultad de Ingeniería – UNSJ

³ Departamento de Informática - F.C.F.M. y N – UNSL

⁴ Instituto de Investigaciones, Facultad de Informática y Diseño, Universidad Champagnat

⁵ Ciencias Exactas Físicas y Naturales - UNLaR

⁶ Alumno Avanzado Licenciatura en Cs. de la Computación - F.C.E.F. y N. - U.N.S.J.

Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan, 0264 4234129

nelson@iinfo.unsj.edu.ar, marite@unsj-cuim.edu.ar, schavez@iinfo.unsj.edu.ar, arianamartinsj@gmail.com, vdiego.unsj@hotmail.com, jnmercado@unsj-cuim.edu.ar, migueljoseguevaratencio@gmail.com, mpiccoli@unsl.edu.ar, mendez-garabettimiguel@uch.edu.ar, rosensteinjavier@uch.edu.ar, monik.gimenez@gmail.com, javiersilleros@gmail.com

Resumen

El surgimiento de Internet de las Cosas y la Computación Móvil junto con la rápida adopción de Cloud Computing y la importancia creciente de big data han cambiado el panorama que presenta la computación distribuida. Sin embargo el volumen de tráfico que se genera desde el borde (o edge) de la red hasta el Cloud donde deberán ser procesados, está llegando a valores elevadísimos que impiden su transferencia. Por otro lado ciertas aplicaciones son sensibles a la latencia como monitoreo de salud, sistemas de tiempo real y sistemas críticos, resultando la demora por la comunicación al cloud, su procesamiento y posterior respuesta, inadmisibles. Dichos datos, procesamiento, control y funciones de red pueden ser realizados en el borde de la red en los dispositivos intermedios. Este modelo de arquitectura es reciente y si bien presenta muchas ventajas, también requiere que se resuelvan muchos problemas. La orquestación de los servicios se refiere a la coordinación y la organización de múltiples servicios, automatizando los procesos al asociar los servicios de diferentes recursos para crear

nuevos servicios compuestos. La misma permitirá mejorar el despliegue y la asignación de los servicios, mantener la continuidad al cloud, mejorar la seguridad y la confiabilidad de los sistemas. Sin embargo cómo realizar esta orquestación y cómo complementar las diversas tecnologías es un problema no resuelto y es el motivo de la presente propuesta de proyecto de investigación.

Palabras clave: *Distributed Computing, Edge Computing, Cloud Computing, Fog Computing, IoT*

Contexto

El presente trabajo se encuadra dentro del área de I/D Procesamiento Distribuido y Paralelo y se enmarca dentro del proyecto de investigación: Orquestación de servicios para la Continuidad de Edge al Cloud, que está en proceso de evaluación para el período 2018-2019. Asimismo el grupo de investigación viene trabajando en proyectos relacionados con la computación distribuida y de alta performance desde hace más de 17 años. En

esta oportunidad se han incorporado investigadores de otras universidades de la región lo cual impactará en todas las actividades planificadas. Las unidades ejecutoras para dicho proyecto son el Departamento e Instituto de Informática de la FCEFyN de la UNSJ.

Introducción

La idea de Cloud Computing fue sugerida por John McCarthy en 1961, sugiriendo utilizar la computación como un servicio público, de forma similar a empresas de servicio o electricidad [1].

Cloud Computing es una tecnología disruptiva que tiene el potencial de mejorar la colaboración, agilidad, escalado y disponibilidad, y ofrece oportunidades para la reducción de costos a través de una computación optimizada y eficiente. El modelo de cloud computing prevé un mundo en el que los componentes se pueden orquestar, aprovisionar, implementar, sacar de servicios, ampliar o disminuir para proporcionar un modelo de asignación al consumo [2].

Esto conlleva una revolución en la forma en que se abstrae y utiliza la infraestructura informática. Algunas de las características que hace que el cloud computing sea deseable incluye; elasticidad, pago por uso y transferencia de riesgo hacia los grandes proveedores de servicios. Por lo tanto, se pueden probar nuevas aplicaciones o ideas con riesgos mínimos, un enfoque que no era factible en la era previa al Cloud [3].

La expansión de Cloud sumado a los requerimientos de los usuarios ha terminado de consolidar este modelo. A la Infraestructura, Plataforma y Software como servicio, se le agregaron: Redes, Almacenamiento, Contenedores, Funciones, Desktop, Base de datos, Seguridad, Recuperación ante desastres, Integración de Cloud, Administración de capital humano, Video y Comunicaciones Unificadas como Servicio entre otras, en definitiva lo que actualmente se conoce como XaaS o Everything a Service [4].

Por otro lado, Internet de las cosas (IoT) es una tendencia reciente de la computación distribuida que integra aspectos de la vida real

a escalas masivas. A pesar de comunicarse con redes inalámbricas, presentan características distintivas: no existen estándares, la variedad de tecnologías es importante, existen varias arquitecturas de software, las velocidades de transferencia en general son bajas y las MTU son menores que las redes IP en general.

En 1999, Kevin Ashton del MIT acuñó el término IoT. En ese momento, las tecnologías de automatización comenzaban a pasar de la fábrica a nuevos entornos como hospitales, bancos y oficinas. A medida que las primeras implementaciones se hicieron cada vez más sofisticadas, las máquinas comenzaron a ser conectadas a otros tipos de dispositivos como servidores, y estos servidores se trasladaron a centro de datos y al Cloud. En la actualidad, IoT puede incluir productos industriales, comerciales, cotidianos (lavavajillas y termostatos) y redes locales de sensores para vigilar granjas y ciudades [5].

Esta soluciones que ofrece IoT promueve la incorporación de los mismos a la red y se pronostica que entre 20 y 50 millones de dispositivos se añadirán a Internet para 2020, creando una economía de más de 3 billones de dólares [6]; en consecuencia, 43 billones de gigabytes de datos serán generados y necesitarán ser procesados en los centros de datos de Cloud. Las aplicaciones que generan datos en dispositivos de usuario, como teléfonos inteligentes y tablets, usan actualmente Cloud como un servidor centralizado, pero pronto se convertirá en un modelo informático insostenible [7].

La importancia de IoT y las capacidades de Cloud, imponen que esta asociación sea por demás necesaria.

Sin embargo, en la actualidad Cloud Computing está encontrando serias dificultades para satisfacer los requerimientos de IoT.

Por ello, surgió como propuesta de solución llevar el almacenamiento, las funciones de red, gran parte del procesamiento hacia el borde de la red, lo que resultó en un nuevo modelo llamado fog computing.

Mist, Fog y Edge Computing

El origen del alcance de edge computing ocurre a fines de la década de 1990, cuando Akamai introduce las redes de entrega de contenido (CDN) para acelerar el rendimiento web. Una CDN utiliza nodos en el borde cerca de los usuarios para captar previamente y almacenar en caché el contenido web. Estos nodos de borde pueden realizar algunas personalizaciones de contenido, como agregar publicidad relevante para la ubicación. Las CDN son valiosas para el contenido de video, por el ahorro de ancho de banda del almacenamiento en caché. Edge computing generaliza y amplía el concepto de CDN al aprovechar la infraestructura cloud [8].

Por otro lado, en un artículo de 2001 se generalizó estos conceptos, introduciéndose el término cyber foraging para la amplificación de las capacidades informáticas de un dispositivo móvil mediante el aprovechamiento de la infraestructura cercana [9].

En 2012, Flavio Bonomi y sus colegas presentaron el término fog computing (informática en la niebla) para referirse a este disperso infraestructura de Cloud [10]. Sin embargo, en este caso, la motivación para la descentralización es la escalabilidad de la infraestructura IoT en lugar del rendimiento interactivo de las aplicaciones móviles.

Los términos Edge y Fog suelen llevar significados diferentes según el autor que los referencie, lo cual genera confusión. Por ello es necesario indicar adecuadamente la definición que se va a utilizar de estos términos.

En particular fog y mist computing, usualmente se asocian a IoT en el extremo, mientras que Mobile Cloud se asocia a Edge Computing para computación móvil.

En la figura 1, se puede visualizar la jerarquía del modelo y las funcionalidades de cada uno de los integrantes (mist, fog, core y center) [11].

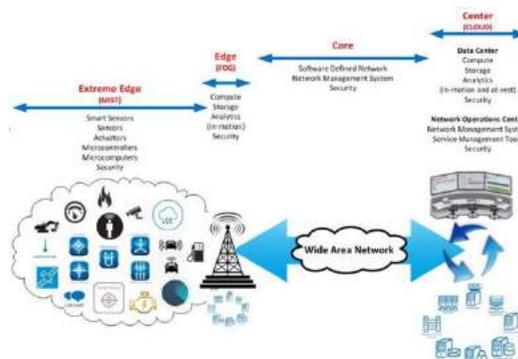


Figura 1. Ubicación de Fog y Mist Computing en la red

Orquestación de servicios para la continuidad

La computación en el extremo (edge) está recién en sus comienzos, tal es así que los investigadores suponen que para su real consolidación se necesitará de la construcción de nuevos estándares y desarrollos.

A pesar de presentar notable ventajas con respecto al modelo centralizado, mantener la continuidad de forma eficiente y efectiva no se puede obtener simplemente agrupando Cloud con el extremo.

Los sistemas del extremo necesitan interactuar con otros, con el cloud y con diversos tipos de usuarios y dispositivos en un ambiente distribuido y heterogéneo.

No solo resulta necesaria una redistribución del cómputo, las funciones de red, el almacenamiento y el monitoreo, sino que además debe ser estratégico para impedir que el sistema colapse o presente demoras inadecuadas. Por otra parte la estrategia debe contemplar si la funcionalidad se llevará al extremo final (mist) o al extremo medio (fog).

Los modelos de programación concurrentes utilizados actualmente no son adecuados para aplicar en la problemática presentada. El modelo de memoria distribuida como MPI no puede ser aplicado dado que se diseñó para otro tipo de arquitecturas y además no puede ejecutarse en el extremo. El modelo de threads solo es aplicable dentro de un mismo equipo. Por lo tanto se deben desarrollar soluciones basadas en el modelo de programación de actor o alguno similar que permita la concurrencia

basado en memoria distribuida pero que sea aplicable al edge. [12].

La seguridad es otro aspecto donde este modelo ofrece respuestas, debido a la demanda de una mayor escalabilidad y seguridad funcional: si las entidades están coordinadas por el Cloud, sigue existiendo el riesgo de que, sin una conexión a Internet confiable, no se pueda garantizar la seguridad funcional [13].

Resultará necesario también procesar eficientemente los recursos, lograr un adecuado balanceo de carga para la distribución entre el borde y el cloud, las API's y la administración y compartición de servicios y comunicaciones en redes definidas por el software y virtualizadas [14].

El problema de la orquestación de la variedad de servicios complejos es un objetivo complicado de lograr por los ambientes altamente dinámicos, la cantidad de aplicaciones para edge instaladas en los dispositivos de usuario final o las cosas y la necesidad de soportar diferentes dominios de aplicación para adaptar el servicio a la heterogeneidad extrema de la infraestructura y para adaptar el servicio al ambiente externo. Por instancia las aplicaciones edge no pueden siempre contar con disponibilidad de los dispositivos de computación más poderosos que puedan ejecutar algoritmos de orquestación complejos. Las condiciones críticas (por ejemplo cuando la infraestructura se cae por desastre climático) la infraestructura debe ser capaz de orquestar los servicios aun en presencia de limitaciones de cómputo con algún grado de resiliencia.

Líneas de Investigación, Desarrollo e Innovación

Las tareas investigación se centrarán en la formulación de la/s hipótesis, que permitan realizar el análisis de los distintos modelos tendientes a lograr la continuidad del Edge al Cloud. Este análisis implica un cuidadoso estudio de las características de cada uno de ellos, con el objetivo de lograr una eficiente orquestación de recursos. Además se utilizarán diversos métodos para la recolección, extracción y clasificación de la

información a partir de grandes volúmenes de datos en base a aspectos tales como dependencia de datos, funciones, sincronizaciones, etc.

Se procurará determinar las características de las soluciones existentes y establecer los puntos a mejorar a través de la computación distribuida. Se formularán la/s hipótesis, para luego analizar y comparar las distintas soluciones que se encuentren, utilizando pruebas experimentales. Una vez generada la propuesta, se validará mediante el diseño e implementación de herramientas (frameworks) distribuidas de alto nivel para su aplicación en la solución de problemas con uso de datos masivos.

Resultados y Objetivos

Resultados Obtenidos

Durante los últimos diez años se trabajó en el área de Computación de Altas Prestaciones, en particular sobre análisis de diversas arquitecturas paralelas y distribuidas, tales como: Cloud Computing (públicos, híbridos y privados), Cluster de commodity, arquitecturas distribuidas de bajo costo y arquitecturas paralelas. Dicha experiencia impulsó esta línea de investigación. El grupo ha realizado publicaciones en el área durante el último año: diez trabajos de investigación en diferentes Congresos y Jornadas, se realizaron tres publicaciones en revistas científicas y se transfirieron los resultados mediante conferencias en eventos científicos.

Se han aprobado ocho tesinas de grado y un trabajo de especialización y se incorporó un becario de investigación categoría alumno.

Objetivos

El objetivo del grupo de investigación es realizar la orquestación de servicios en entornos de continuidad Edge al Cloud, analizando distintas estrategias y evaluando cómo se comportan los parámetros de desempeño comparado con el modelo de cloud centralizado.

Formación de Recursos Humanos

El equipo de trabajo está compuesto por los once docentes-investigadores que figuran en este trabajo de las universidades de San Luis, Champagnat, La Rioja y San Juan, un becario de iniciación de CONICET y ocho alumnos.

Se están realizando dos tesis doctorales, sobre procesamiento de datos masivos una en entornos distribuidos y otra en entornos paralelos. Se espera realizar también una tesis de maestría sobre control topológico para reducción de interferencia en redes IoT. Se están realizando cinco tesinas de licenciatura una sobre evaluación de algoritmos de algebra lineal sobre arquitecturas diversas, otra sobre Cloud Computing Privado, otra sobre arquitecturas de seguridad sobre Cloud Privado. Además se espera aumentar el número de publicaciones. Por otro lado también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación y asesoramiento a empresas y otros organismos del estado.

Referencias

- [1] V Rajaraman (Indian Institute of Science) (2014), "Cloud Computing", <http://www.ias.ac.in/article/fulltext/reso/019/03/0242-0258>
- [2] Cloud security alliance (2011), "Security guidance for critical areas of focus in cloud computing v3.0", <https://downloads.cloudsecurityalliance.org/initiatives/guidance/csaguide.v3.0.pdf>.
- [3] Niko Mäkitalo et al, (2018), "Safe, Secure Executions at the Network Edge Coordinating Cloud, Edge, and Fog Computing", IEEE Software, 0740-7459/18 IEEE.
- [4] CBS Interactive Inc (2017), "Special Report: The future of Everything as a Service".
- [5] Joe Biron and Jonathan Follett (2016), "Foundational Elements of an IoT Solution the

Edge, the Cloud, and Application Development", O'Reilly Media, Inc.

- [6] Garnet: "Gartner Says 6.4 Billion Connected Things Will Be in Use in 2016, Up 30 Percent From 2015".

<http://www.gartner.com/newsroom/id/3165317>.

- [7] Blesson Varghese, Nan Wang, Dimitrios S. Nikolopoulos, Rajkumar Buyya, "Feasibility of Fog Computing", arXiv:1701.05451v1cs.DC

- [8] Dilley J. et al. (2002), "Globally Distributed Content Delivery," IEEE Internet Computing, vol. 6, No.5, 2002, pp.50–58.

- [9] Satyanarayanan M. (2001), "Pervasive Computing: Vision and Challenges," IEEE Personal Comm., vol. 8 No.4, 2001, pp.10–17

- [10] Bonomi F. et al (2012), "Fog Computing and Its Role in the Internet of Things", Proc. 1st Edition MCC Workshop Mobile Cloud Computing (MCC12), 2012, pp.13–15.

- [11] Martin M., GTS Network Services Group - IBM (2016), "Edge Computing - Where data comes alive!" <https://www.linkedin.com/pulse/edge-computing-where-data-comes-alive-michael-j->

- [12] Hewitt C., Bishop P. and Steiger R. (1973), "A Universal Modular ACTOR Formalism for Artificial Intelligence". Proceedings of the 3rd international joint conference on Artificial.

- [13] Shivam Sharma (2018), "Cloud computing security challenges and solutions", International Research Journal of Computer Science (IRJCS). Issue 02, Volume 5.

- [14] Vahid Dastjerdi A. and Buyya R. (2016) "Fog Computing: Helping the Internet of Things Realize Its Potential", Computer, IEEE Computer Society - 0018-91 62/16 /2016.

Modelado y Simulación de Sistemas de Gran Escala

Rubén Apolloni¹, Alicia Castro¹, Verónica Gil-Costa², Fernando Loor², Cristian Tisera¹
and Guillermo Trabes¹

⁽¹⁾Departamento de Informática
Facultad de Ciencias Físico
Matemáticas y Naturales
Universidad Nacional de San Luis
Ejército de los Andes 950, 1° piso.
(02652-420823)

⁽²⁾CCT-XONICET San Luis
Almirante Brown 907
(02664-421654)

Resumen

Las simulaciones computacionales permiten entender los elementos y patrones que pueden alterar un sistema y puede ser utilizada para estudiar sistemas complejos, incluyendo aquellos que son analíticamente intratables. La simulación también es utilizada cuando no se tienen los recursos computacionales disponibles. En estos casos se puede estimar diferentes parámetros como la escalabilidad de los sistemas, a medida que se simula el aumento del número de recursos asignados. En otras palabras, utilizando técnicas de modelado y simulaciones eficientes es posible diseñar, desarrollar y evaluar sistemas de gran escala. En este trabajo, se presentan los objetivos, trabajo realizado y desafíos que aborda el grupo de investigación de la Universidad Nacional de San Luis, para abordar los temas que involucra el diseño de simulaciones de sistemas complejos y de gran escala para procesar grandes volúmenes de datos e información.

Palabras clave: *Sistemas Complejos y de Gran Escala, Simulación, Modelado.*

1. Contexto

La línea de investigación que se presenta en este trabajo “Modelado y Simulación de Sistemas de Gran Escala” pertenece al proyecto PROICO “Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos” de la UNSL. Esta línea de

investigación, se basa en el uso de técnicas de simulaciones eficientes para diseñar y desarrollar sistemas de gran escala como lo son los sistemas de búsqueda Web y sistemas de evaluación y soporte a las decisiones para situaciones de desastres naturales. Con estas simulaciones se pueden probar alternativas en el diseño y predecir su impacto sobre un sistema real por medio de métricas de rendimiento, lo cual permite determinar el costo-beneficio e implementar la aplicación en un hardware real.

Existen diferentes técnicas de simulación como la Simulación Aproximada [Marin13], Redes de Petri [Petri62] [Peterson77], simuladores orientados a procesos, simulación paralela, entre otros. En particular, las investigaciones abordadas se enfocan en considerar el modelado de aplicaciones complejas y de gran escala donde intervienen diferentes variables de incertidumbre para plataformas paralelas y aplicadas a sistemas, que no pueden ser probadas en sistemas y hardware reales sin acarrear un costo económico que puede ser de gran impacto en el caso de que las nuevas aplicaciones o los algoritmos no cumplan con sus objetivos.

2. Introducción y Motivación

El desarrollo e implementación de sistemas de gran escala es una tarea compleja y más aún cuando se espera que sean capaces de seguir funcionando en situaciones de desastre natural. Varios son

los factores a considerar bajo esta premisa, como por ejemplo: el tiempo de desarrollo, arquitectura de software, el comportamiento dinámico de los usuarios, cortes de luz, destrucción de infraestructura de telecomunicación, etc., lo cual finalmente se traducen en una costosa inversión, que de no llevarla a cabo correctamente podría poner vidas en riesgo. Para disminuir dichos riesgos se hace necesario contar, de ante mano, con herramientas que permitan modelar y analizar las diversas situaciones que se podrían producir al usar el sistema, libre de riesgos y en un ambiente seguro. Es en este punto donde se destaca el uso de la simulación computacional, una disciplina ampliamente estudiada cuando el análisis de un sistema complejo ya no puede ser llevado a cabo mediante métodos analíticos.

La Simulación de Eventos Discretos, desde ahora DES (en inglés Discrete Event Simulation), se utiliza para representar sistemas complejos como una secuencia ordenada de eventos y donde cada evento posee el time-stamp del sistema físico representado. Cuando dichos sistemas son muy costosos en términos de procesamiento, disco y uso de memoria, su simulación también puede ser costosa, siendo éste el caso, se puede decir que el resultado es una simulación de gran escala que suele enfrentarse mediante el uso de Simulación Paralela de Eventos Discretos, desde ahora PDES (en inglés Paralell Discrete Event Simulation). Este tipo de simulación busca aprovechar al máximo las capacidades de los clusters de cómputo implementando diversas técnicas, por ejemplo: la distribución y balance de carga de trabajos en las unidades de cómputo y sincronización de eventos concurrentes mediante los protocolos conservativos (evita violar las condiciones de causalidad) y optimistas (avanza permitiendo violar las condiciones de causalidad y realiza roll-back cuando esto sucede).

El objetivo principal, de esta línea de investigación, es investigar el uso de simuladores eficiente y eficaces para evaluar el rendimiento de nuevos

algoritmos paralelos y distribuidos para aplicaciones de datos masivos. Un punto crucial es probar algoritmos y estrategias que permitan balancear la carga de trabajo de los componentes a evaluar. En particular, se propone (a) Investigar el enfoque forward-in-time para simulaciones de evoluciones de poblaciones. (b) Obtener simuladores que sean escalables, reduzcan el consumo de memoria y entreguen resultados en el menor tiempo posible. (c) Investigar la aplicación de BSP y Multi-BSP para modelar problemas de gran escala y estimar los tiempos de ejecución utilizando los parámetros provistos por los modelos. (d) Continuar estudiando y desarrollado simuladores basados en autómatas celulares para situaciones de evacuaciones. (e) Continuar aplicando el conocimiento adquirido en esta línea para realizar actividades de servicio y poner en práctica los nuevos algoritmos/enfoques desarrollados.

A continuación, se describen algunas de las técnicas, herramientas y plataformas utilizadas.

3. Investigación y Desarrollo

3.1 Capacity Planning

Capacity planning [Klosterboer11] es un proceso iterativo, que explota el comportamiento histórico de la gente para predecir el crecimiento de las necesidades de éstos a través del tiempo y prever futuros requerimientos de hardware en alzas inesperados de peticiones. La empresa Amazon cuenta con metodologías de capacity planning privadas, sin embargo sólo se sabe que para poder predecir el crecimiento de las necesidades de sus centros de datos actualmente utilizan Aprendizaje Máquina (Machine Learning), [geekwire17].

De manera similar, Google utiliza Aprendizaje Máquina para hacer un capacity planning enfocándose en optimizar el Uso de Energía Eficiente (Power Usage Efficient, PUE) [Gao14].

En algunos trabajos previos, se han desarrollado modelos y simuladores para analizar la cantidad de recursos que deben asignarse para un motor de búsqueda Web. En particular, en [Gil12] se propone utilizar un modelo de Redes de Petri.

Posteriormente, en [Gil13] se evalúan diferentes algoritmos de despliegues para mejorar el rendimiento del sistema.

3.2 Simulación Paralela Aproximada

La sincronización de eventos en PDES suele llevarse a cabo mediante estrategias conservativas y optimista. Sin embargo, ambas estrategias suelen ser costosas en términos computacionales y en tiempo de ejecución. Otra estrategia válida es la propuesta en [Marin13], donde se ha expuesto un nuevo enfoque mediante una estrategia de simulación Aproximada, la cual plantea realizar una sincronización optimista que permite ejecutar eventos con time-stamp que violan la condición de causalidad. Es decir, ejecutar eventos con time-stamp menor al tiempo global de la simulación, esto sin realizar roll-back. Su lógica se basa en ir calculando periódicamente barreras tiempo dinámicas, gracias a una técnica de ventanas deslizantes, que posibilitan lo anterior. Con este método lograron demostrar una eficiente paralelización de los eventos.

Los modelos de simulación proveen una mejora en términos de tiempo de ejecución. Bajo este paradigma el modelo de simulación es particionado en diferentes procesadores físicos que se comunican entre sí intercambiando mensajes (eventos).

En particular, en el trabajo presentado en [Marin13] se propone priorizar los tiempos de ejecución y perder hasta un 5% de precisión en los resultados finales de la simulación. Este trabajo se basa en la utilización del modelo de computación paralelo BSP [Val90], donde el cómputo se organiza en una secuencia de supersteps. Durante un superstep, cada

procesador realiza operaciones sobre sus datos locales y/o envía mensajes a otros procesadores. Al final de cada superstep se ejecuta una barrera de sincronización que permite que los mensajes enviados en el superstep actual, se encuentren disponibles en las colas de entrada de los procesadores destinos en el próximo superstep.

3.3 Modelos de Simulación

Los modelos de computación se han usado como un puente entre algoritmos paralelos y arquitecturas de hardware. El modelo Bulk-Synchronous Parallel (BSP) [Val90] es un modelo muy conocido originalmente diseñado para ejecutar algoritmos distribuidos en clusters de procesadores de un solo núcleo. El modelo Multi-BSP, que amplía el modelo BSP clásico, fue propuesto recientemente para procesadores multi-core.

Sin embargo, el modelo Multi-BSP presenta algunas restricciones tales como las sincronizaciones explícitas entre los cores, presentando algunos desafíos que se deben tener en cuenta para modelar adecuadamente los algoritmos paralelos. Por lo tanto, es importante explorar la utilización y el alcance de estos modelos para las arquitecturas multi-core [Trabes18].

3.4 Simulación de Evacuaciones

En los últimos años se ha incrementado el interés en desarrollo de modelos de procesos de evacuaciones de emergencias. Nuestra propuesta, se enfoca en el estudio de evacuación de edificaciones, el cual radica en que la totalidad de sus ocupantes, en cualquier instante, deben tener la posibilidad de desplazarse hasta un lugar seguro en el tiempo adecuado con las suficientes garantías de seguridad.

Debido a la dificultad en cómputo para resolver este tipo de problemas, un modelo de simulación complejo puede requerir de modernas plataformas computacionales de alta performance para acelerar a gran escala las simulaciones de

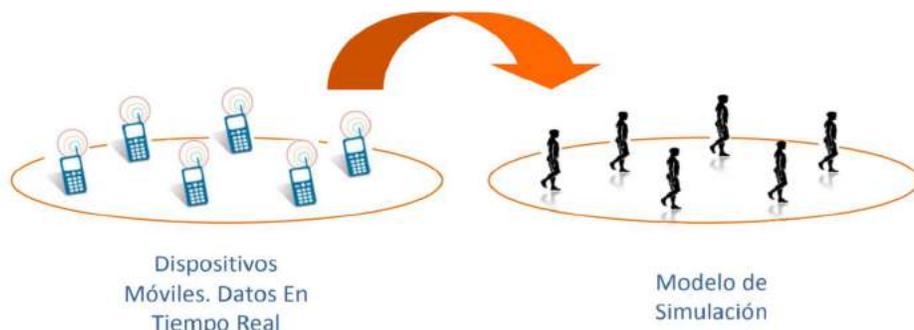


Figura 1: Datos en tiempo real enviados a un modelo de simulación

eventos discretos y promover la reutilización y la interoperabilidad de los componentes de simulación.

Actualmente se cuenta con la implementación de un modelo de simulación que permite simular el comportamiento de un conjunto de individuos en un ambiente cerrado con un número fijo de salidas. El principal objetivo de este modelo, es el de recrear la situación en la cual un conjunto de personas debe desalojar un recinto debido a la amenaza de fuego. En este modelo es factible obtener la información necesaria a través de sensores o dispositivos móviles, que será utilizada como entrada para nuestro modelo (Fig. 1).

También se han desarrollado modelos de simulación híbrida [TPL12] que consiste en dos sub-modelos denominados ambiental (EsM) y peatón [PsM]. Este modelo junto con la metodología computacional permite construir ambientes artificiales poblados con agentes autónomos que son capaces de interactuar entre sí.

4. Resultados

La línea de investigación y las técnicas descritas en la sección anterior involucra una serie de desarrollos individuales que en su conjunto logran obtener el objetivo planteado. Este objetivo contempla el modelado y diseño de sistemas complejos mediante diferentes herramientas, algunas de las cuales han sido desarrolladas por los integrantes del proyecto.

Mediante estas herramientas se pretende facilitar el entendimiento del sistema, realizar profiling a la ejecución de los algoritmos y mejorar el análisis de resultados.

Los resultados obtenidos hasta el momento son:

- Evaluar el beneficio de la colaboración entre un motor de búsqueda Web y un proveedor de servicio de internet (ISP) [Ros14]. En particular, en este trabajo se propone utilizar una jerarquía de memorias cachés ubicadas en los nodos de las redes P2P y en el servidor ISP, con el objetivo de reducir el tráfico de consultas dirigidas al motor de búsqueda. A cambio, el motor de búsqueda puede reducir los costos por acceder a él y entregar información más precisa respecto del tiempo de vida de los resultados.
- En [Hid14] y [Gil14b] se realiza una evaluación del consumo de energía requerido tanto por la red de comunicación interna al ISP y la red de comunicación entre el ISP y el Motor de búsqueda.

Los resultados esperados son:

- Generar un modelo de las aplicaciones de gran escala. Modelar los componentes más relevantes y la comunicación entre los mismos.
- Desarrollar simuladores secuenciales de las aplicaciones de gran escala. Validar los resultados de los simuladores y verificar que el comportamiento (flujo de datos y operaciones) sea correcto.

- Proponer enfoques paralelos eficientes de los simuladores desarrollados. Utilizar estos simuladores para realizar estudios de capacity planning.
- Transferencia de logros obtenidos.

5. Formación de Recursos Humanos

Actualmente, se cuenta con tres doctores en ciencias de la computación realizando la investigación y dirección de los doctorandos involucrados en la línea. Se cuenta con tres estudiantes de doctorando. Dos de ellos se encuentran en su segundo año de trabajo. El tercer doctorando se encuentra en su primer año, de iniciación a la investigación. El primer estudiante realiza su tesis en el desarrollo y estudio de modelos para la automatización del despliegue de aplicaciones en plataformas paralelas. El segundo estudiante, realiza su tesis doctoral en el estudio y desarrollo de plataformas escalables para casos de desastres naturales utilizando técnicas de crowdsourcing.

Mediante este trabajo de investigación se podrán formar profesionales que puedan modelar, diseñar e implementar algoritmos eficientes (previamente evaluados y analizados con un mínimo porcentaje de error mediante las técnicas de modelado y simulación) que se ejecuten en sistemas complejos donde intervienen una gran cantidad de variables y requieren el procesamiento de datos masivos.

6. Bibliografía

- [Gil12] V. Gil-Costa, J. Lobos, A. Inostrosa-Psijas and M. Marin. "Capacity Planning for Vertical Search Engines: An approach based on Coloured Petri Nets" Petri Nets 2012.
- [Gil13] A. Inostrosa-Psijas, V. Gil-Costa, M. Marin and E. Feuestein. "Service Deployment Algorithms for Vertical Search Engines". PDP 2013.
- [Gil14b] V. Gil-Costa, E. Rosas, N. Hidalgo and M. Marin. "Simulation Model for energy assessment in ISP and Web Search Engine collaboration". Summer Simulation Multi-Conference, pp. 23:1--23:8, 2014. [Geekwire17] <https://www.geekwire.com/2017/amazon-web-services-uses-machine-learning-make-capacity-planning-decisions/>, at Pacific Science Center's 14th Annual Foundations of Science Breakfast
- [Gao14] Gao, J., Jamidar, R. (2014). Machine learning applications for data center optimization. Google White Paper
- [Hid14] "Assessing energy efficiency in ISP and Web Search Engine collaboration". N. Hidalgo, E. Rosas, V. Gil-Costa and M. Marin. FINA. pp. 299-304. May 13 - 16, 2014.
- [Klosterboer 11] Klosterboer, L. (2011). ITIL capacity management. Upper Saddle River, NJ: IBM Press. [Google Scholar]
- [Marin13] M. Marin, V. Gil-Costa, C. Bonacic and R. Solar. pproximate Parallel Simulation of Web Search Engines. PADS 2013
- [Peterson77] J.L. Peterson. Petri nets. Computing Surveys, 9:223-252, 1977.
- [Petri62] C.A. Petri. "Communication with Automata" New York: Griffiss Air Force Base. Tech. Rep. RADC-TR-65-377, vol.1, Suppl. 1, 1962.
- [Ros14] E. Rosas, N. Hidalgo, M. Marin, V. Gil-Costa. Web Search Results Caching Service for Structured P2P Networks. Future Generation Computer Systems (Elsevier). Volume 30, Pages 254-264, January 2014.
- [TPL12] P. Tisera A. M. Printista, and E. Luque. A Hybrid Simulation Model to Test Behaviour Designs in an Emergency Evacuation. ICCS, (2012)
- [Trabes18] G. Trabes, V. Gil-Costa, M. Printista and M. Marin. Multi-BSP vs. BSP: A Case of Study for Dell AMD Multicores. PDP 2018-
- [Val90] L. G. Valiant. A bridging model for parallel computation. Comm. ACM, 33(8) 1990.

Procesamiento Computacional Paralelo con Metaheurísticas Híbridas para la Reducción de Incertidumbre en Modelos de Incendios Forestales

Bianchini Germán¹, Caymes-Scutari Paola^{1,2}, Méndez-Garabetti Miguel^{1,2},
Tardivo María Laura^{1,2,3}

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido
Departamento de Ingeniería en Sistemas de Información
Facultad Regional Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

³Departamento de Computación, Universidad Nacional de Río Cuarto, Río Cuarto,
Córdoba, Argentina

gbianchini@frm.utn.edu.ar, pcaymesscutari@frm.utn.edu.ar, mmendez@mendoza-conicet.gob.ar,
lauratardivo@dc.exa.unrc.edu.ar

RESUMEN

La predicción del comportamiento de incendios forestales no es una tarea sencilla ya que dicho proceso se ve afectado por la falta de precisión o incertidumbre en los parámetros de entrada. En base a esto, resulta importante desarrollar métodos que permitan tratar la incertidumbre posibilitando la obtención de predicciones más precisas y confiables. En el proyecto que aquí se expone se propone el desarrollo de un método de reducción de incertidumbre denominado Sistema Estadístico Evolutivo Híbrido con Modelo de Islas (HESSIM). En HESSIM se plantea un método que combine la fuerza de tres metaheurísticas poblacionales evolutivas: Algoritmos Evolutivos (EA), Evolución Diferencial (DE) y Optimización por Cúmulo de Partículas (PSO), bajo un esquema de combinación colaborativa basado en migración mediante modelo de islas y HPC. HESSIM corresponde a la continuación de las líneas abiertas

resultantes de proyectos previos en los cuales se implementaron dos versiones de ESS-IM: la primera con Algoritmos Evolutivos, y la segunda con Evolución Diferencial. De este modo surge la idea de desarrollar una versión híbrida en un único método, en vista del potencial aportado por cada una de las metaheurísticas poblacionales en forma aislada, y añadiendo además una tercera (PSO). Dado que las arquitecturas paralelas se han convertido en una herramienta importante en muchos campos de la ciencia por los beneficios que aporta a la hora de efectuar los cálculos, y debido a la naturaleza intrínsecamente paralela de las tres metaheurísticas elegidas, HESSIM se implementará siguiendo un esquema de procesamiento paralelo.

Palabras clave: Evolución Diferencial, Algoritmos Evolutivos, Optimización por Cúmulo de Partículas, HPC, Reducción de Incertidumbre.

CONTEXTO

El proyecto se encuentra en ejecución desde enero del presente año 2018, llevándose a cabo en el marco del LICPaD (Laboratorio de Investigación en Cómputo Paralelo/Distribuido) dentro del ámbito de la UTN-FRM. Cuenta con el financiamiento de la UTN a través del proyecto SIUTIME0004736TC y mantiene la línea de proyectos anteriores en los cuales se desarrolló inicialmente el método ESS (*Evolutionary-Statistical System*) [1], el cual fue evolucionando, pasando por diversas versiones incrementales que se comentan en la siguiente sección, hasta llegar a la actual propuesta.

1. INTRODUCCIÓN

Las nociones iniciales que anteceden al presente proyecto provienen, en primer lugar, del método S^2F^2M (Sistema Estadístico para la Gestión de Incendios Forestales) [2, 3], el cual utiliza análisis estadístico y paralelismo. Este método realiza un elevado número de simulaciones sobre un conjunto de diferentes configuraciones de parámetros de entrada (tipo de vegetación, humedad, velocidad del viento, pendiente del terreno, etc.) denominados escenarios. Todos los escenarios posibles son generados en forma discreta considerando un cierto dominio a través de un experimento factorial [4] y el modelo es evaluado con cada conjunto de valores. Los resultados se combinan para determinar la tendencia en el comportamiento del modelo, ajustándolo con la observación actual del mismo. El patrón hallado es entonces considerado para la predicción del siguiente paso. Posteriormente a S^2F^2M se desarrolló el método ESS (Sistema Estadístico Evolutivo) [1], el cual corresponde a una

mejora de S^2F^2M lograda mediante la introducción de Algoritmos Evolutivos (EA) [5], en su modalidad paralela (PEA) [6] (con un esquema de Única Población y Evaluación en Paralelo), para tratar de forma más eficiente la generación de escenarios (en ESS el concepto de “conjunto de escenarios” es reemplazado por el de “población de individuos”) donde un escenario particular define a un individuo. ESS, al igual que S^2F^2M , ha sido implementado en un entorno paralelo/distribuido. Seguidamente se realizó una variante de ESS en la que se incrementó el nivel de paralelismo del PEA con un esquema de Múltiples Poblaciones y Migración [6]. Esta implementación se denominó Sistema Estadístico Evolutivo con Modelo de Islas (ESS-IM, por sus siglas en inglés) [7]. ESS-IM ha permitido obtener mejores resultados que sus antecesores mediante el incremento de diversidad de individuos en el PEA, logrado mediante la iteración e interacción de múltiples poblaciones de individuos o islas.

Si bien los resultados obtenidos con ESS-IM han sido superiores a los de ESS y S^2F^2M [X], en este proyecto proponemos el desarrollo de una nueva versión de ESS-IM, donde la metaheurística central del sistema (en una versión EA y en otra DE) sea reemplazada por un modelo híbrido que combine el funcionamiento de varias metaheurísticas en un esquema colaborativo. Por ello es que hemos evaluado diversas opciones que permitiesen la mejora de los métodos actuales y el desarrollo de otras metodologías para resolver el problema de la incertidumbre de maneras más eficientes. En esta primera aproximación hemos elegido las metaheurísticas poblacionales a) Optimización por Cúmulo de Partículas (PSO) [9, 10], b) Algoritmos Evolutivos (EA) [5] y c) Evolución Diferencial [11], dadas sus

características y posibilidades de ser utilizados en el problema que nos concierne. Es importante mencionar que, si bien el método que se propone se desarrollará en el marco del presente proyecto, previamente hemos probado una versión preliminar, la cual implementamos sólo con Algoritmos Evolutivos y Evolución Diferencial. Dicha versión ha obtenido resultados alentadores, ya que ha superado los resultados de las metodologías previamente desarrolladas, sin considerar que aún resta realizar estudios más exhaustivos y concluir el esquema que incorpore a PSO.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

La predicción del comportamiento de los incendios forestales es considerada una herramienta muy útil, ya que permite determinar las zonas que serán afectadas por el fuego, ayudando a tomar medidas de evacuación y de gestión de recursos de forma eficiente. La predicción de cualquier fenómeno natural es una tarea compleja, tanto por el diseño de los modelos involucrados como así también por la presencia de incertidumbre en la información con la que dicho modelo es alimentado. En base a esto, la predicción de un incendio forestal no escapa a esta situación, ya que su comportamiento está determinado por una serie de parámetros que usualmente no pueden ser medidos en tiempo real. Esta falta de precisión en los parámetros de entrada impacta directamente en la salida del modelo, impidiendo a las técnicas de predicción clásicas [12, 13, 14] obtener predicciones aceptables. En este contexto, se debe recurrir a métodos de reducción de incertidumbre que permitan generar predicciones aceptables independiente-

mente de la imprecisión en los parámetros de entrada.

En síntesis, la problemática existente a raíz de la falta de exactitud presente en los parámetros de entrada en cualquier modelo científico o físico, puede producir consecuencias dramáticas en la salida del mismo particularmente si se trata éste de un sistema crítico. Por tal razón, la presente línea de investigación consiste en el desarrollo de métodos computacionales (y en este proyecto en concreto el desarrollo de HESSIM), que se enfoquen en el tratamiento de la incertidumbre de los valores de entrada de modelos para lograr así una predicción lo más confiable posible por parte del mismo. El método propuesto en este caso utilizará como técnica de optimización una metaheurística híbrida basada en Algoritmos Evolutivos, Evolución Diferencial y Optimización por Cúmulo de Partículas, bajo un esquema de integración colaborativa. Es importante mencionar que HESSIM operará con un esquema de paralelización basado en islas con doble jerarquía master-worker, lo que le permitirá implementar las diferentes metaheurísticas con múltiples poblaciones y migración. Se supone que tal esquema ayudará a conseguir un mejor y mayor aprovechamiento del paralelismo potencial de tareas presente en el método.

3. RESULTADOS ESPERADOS

En primer lugar, se espera obtener mejoras en la calidad de la predicción alcanzada por el método, a raíz del incremento en la diversidad de casos lograda mediante la multiplicidad de poblaciones y la heterogeneidad provista por las distintas metaheurísticas involucradas. En segundo lugar, mejorar los tiempos de ejecución de la metodología. Para verificar tales mejoras, los resultados de HESSIM han de ser

comparados con los producidos por las versiones anteriores (ESS-IM, ESS y S²F²M) a través de una serie de experimentos que involucren la utilización de datos de incendios reales de diversas zonas forestales. Hasta el momento, los métodos han sido comparados a través de casos correspondientes a quemas controladas desarrolladas en el campo para tal fin.

Por otra parte, el proyecto está enmarcado en una línea de investigación a más largo plazo, lo que asegura el avance científico por el interés de la propuesta y el tema planteado, lo cual brinda un amplio y prometedor campo de trabajo.

4. FORMACIÓN DE RECURSOS HUMANOS

Además de la innovación científica, se espera que el presente proyecto provea de medios económicos para el fortalecimiento de las líneas de investigación relacionadas y brinde soporte a las actividades que se desprendan del mismo. Tales líneas se extenderán mediante la incorporación de nuevos becarios y estudiantes avanzados, principalmente de la carrera de Ingeniería en Sistemas de Información, aunque no de forma excluyente, facilitando y propiciando de esta manera el acercamiento de los nuevos profesionales al área del HPC. Por su parte, a nivel de postgrado, y en particular referido a estudiantes de doctorado, esta línea de investigación cuenta con dos tesis de doctorado en curso que se espera estén concluyendo durante el primer año de proyecto: la primera, perteneciente al Ing. Miguel Méndez Garabetti, cuyo plan de tesis doctoral se relaciona con este proyecto, y la segunda, a cargo de la Lic. María Laura Tardivo, cuya temática también se vincula estrechamente, dado

que se especializa en distintos métodos y posibilidades referentes a DE. Ambos cursan el doctorado en Ciencias de la Computación de la Universidad Nacional de San Luis, y son becarios de CONICET.

5. BIBLIOGRAFÍA

- [1] Bianchini, G., Caymes Scutari, P., Méndez Garabetti, M., Evolutionary-Statistical System: a Parallel Method for Improving Forest Fire Spread Prediction, *Journal of Computational Science (JOCS, Elsevier)*. Vol. 6, pp. 58-66, 2015.
DOI: 10.1016/j.jocs.2014.12.001
- [2] Bianchini, G., Denham, M., Cortés, A., Margalef, T., & Luque, E., Wildland fire growth prediction method based on Multiple Overlapping Solution. *Journal of Computational Science*, 1(4), 229–237, 2010.
DOI: 10.1016/j.jocs.2010.07.005
- [3] Bianchini, G., Denham, M., Cortés, M., Margalef, T., Luque, E., Improving forest-fire prediction by applying a statistical approach, *Forest Ecology and Management*, 234 (supplement 1), p. S210, 2006.
- [4] Montgomery D.C, Runger G.C.: *Applied Statistics and Probability for Engineers*. 6th Edition. Limusa Wiley & Sons, New Jersey, 2014.
- [5] Goldberg, D.E., Genetic and evolutionary algorithms, *Come of age*, *Communications of the ACM*, 37 (3) pp. 113119, 1994.

- [6] Cantú Paz E., A survey of Parallel Genetic Algorithms. *Calculateurs Parallèles, Réseaux et Systems Repartis*, 10(2), pp. 141-171, 1998.
- [7] Méndez-Garabetti, M., Bianchini, G., Tardivo, M. L., & Caymes-Scutari, P., Comparative analysis of performance and quality of prediction between ESS and ESS-IM. *Electronic Notes in Theoretical Computer Science*, 314, 45–60, 2015. DOI: 10.1016/j.entcs.2015.05.004
- [8] Tardivo, M.L., Caymes Scutari, P., Bianchini, G., Méndez Garabetti, M., Cencerrado, A., Cortés, A., A comparative study of evolutionary statistical methods for uncertainty reduction in forest fire propagation prediction. *Procedia Computer Science*, 108, 2018–2027, 2017. DOI: 10.1016/j.procs.2017.05.252
- [9] Engelbrecht, A., *Fundamentals of Computational Swarm Intelligence*. John Wiley & Sons, Ltd., 2005.
- [10] Kennedy, J., Eberhart, R., Particle Swarm Optimization. *Proceedings of the IEEE International Joint Conference on Neural Networks*, pp. 1942-1948, 1995.
- [11] Méndez-Garabetti, M., Bianchini, G., Tardivo, M.L., Caymes-Scutari, P., Análisis de Metaheurísticas Poblacionales para su Aplicación en Métodos de Reducción de Incertidumbre. Séptimo Encuentro de Investigadores y Docentes de Ingeniería EnIDI 2013, Los Reyunos, San Rafael. Mendoza, Argentina, pp. 543-557, 2013.
- [12] Wallace, G., A Numerical Fire Simulation-Model. *International Journal of Wildland Fire*, 3(2), 111, 1993. DOI: 10.1071/WF9930111
- [13] Finney, M. A., & Finney, M. A. (n.d.). FARSITE: Fire Area Simulator-model development and evaluation.
- [14] Lopes, A. M. G., Cruz, M. G., Viegas, D. X., & Lopes, A. M. G. (n.d.). FireStation - An integrated software system for the numerical simulation of fire spread on complex topography. Retrieved from <http://adai.dem.uc.pt>

Performance de arquitecturas multiprocesador: técnicas de modelado y simulación en HPC y Cloud Computing.

Diego Encinas^{1,2}, Jimena Jara¹, David Rosatto¹, Román Bond¹, Maximiliano Belizán¹,
Martin Morales^{1,3}

¹Instituto de Ingeniería y Agronomía - Universidad Nacional Arturo Jauretche

²Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática -
UNLP – Centro Asociado CIC

³ Unidad CodApli - Facultad Regional La Plata - UTN

dencinas@unaj.edu.ar, elchejime@gmail.com, danielrosatto@gmail.com,
roman.alejandro.b@gmail.com, maximiliano.h.belizan@gmail.com,
martin.morales@unaj.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio de la performance de las arquitecturas multiprocesador y Cloud Computing a través de modelos de simulación. Enfocando a la obtención de herramientas que permitan predecir la eficiencia del sistema ante posibles escenarios y reconfigurar el sistema físico. Analizando los diferentes componentes del sistema que pueden influir en las prestaciones significativamente y pueden llegar a modelarse y/o reconfigurarse.

Palabras clave: *Arquitecturas Multiprocesador. Simulación. Sistema E/S paralela. Modelado y Simulación basado en agentes (Agent-Based Modeling and Simulation, ABMS). Cloud Computing. CloudSim.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación

“Modelado y Simulación en Cómputo de Altas Prestaciones (HPC). Aplicaciones en arquitecturas multiprocesador, sistemas paralelos y redes de datos” de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 186/15. Además, el Proyecto integra el Programa “Tecnologías de la información y la comunicación (TIC) en aplicaciones de interés social” de la UNAJ.

En el tema existe un convenio de colaboración en actividades de Investigación y Postgrado con el Instituto de Investigación en Informática – LIDI de la Universidad Nacional de La Plata. Además, se ha iniciado una colaboración en actividades de investigación con el Área de Computación del Instituto de Ciencias de la Universidad Nacional de General Sarmiento.

Por otra parte, se tiene financiamiento en el marco del programa “Universidad, Diseño y Desarrollo Productivo” del Ministerio de Educación a través de varios proyectos aprobados en la UNAJ.

Introducción

El crecimiento sostenido en la demanda del poder de cómputo remarca la necesidad de sistemas con enfoques de paralelización masiva y cómputo de alta performance (HPC, High Performance Computing) [1]. Los clusters se han convertido en uno de los enfoques principales para lograr paralelismo a bajo costo. Una noción extendida lo constituye la utilización de grid computing y más recientemente cloud computing. Independientemente de la solución, estos sistemas constan de un gran número de componentes incluyendo nodos de procesamiento, bancos de memoria, discos, entre otros.

En cuanto a las herramientas de simulación, CloudSim es un framework desarrollado en Java que provee las APIs necesarias para que el usuario genere una simulación de un Data Center funcionando como servidor de nube capaz de simular la ejecución de CloudLets.

Sistemas de E/S Paralela

Las exigencias en los sistemas de E/S paralelos se han incrementado debido al aumento en número, velocidad y potencia de las unidades de procesamiento en los clusters. También las aplicaciones científicas que utilizan cómputo de altas prestaciones acrecientan estos requerimientos.

En muchos casos, el cuello de botella de los sistemas paralelos es la E/S de estos sistemas dada las exigencias que debe afrontar [2]. La E/S Paralela es esencial para emparejar el avance de las arquitecturas de los procesadores y el rápido crecimiento de la capacidad computacional. Aunque la arquitectura jerárquica de memoria multinivel puede evitar grandes pérdidas de prestaciones debido a los retardos de acceso a disco, la

capacidad de memoria es limitada. Además, como la capacidad computacional aumentará, la disponibilidad de memoria por core decrecerá, especialmente si la escala de los sistemas de HPC se proyecta a millones de cores o más. Varias simulaciones científicas y de ingeniería de áreas críticas de investigación, tales como la nanotecnología, astrofísica, clima y energía física están convirtiéndose en aplicaciones intensivas de datos. Para poder disminuir la brecha entre CPUs-E/S se deben identificar los factores que influyen en las prestaciones y proponer nuevas soluciones [3] [4].

En el área de tolerancia a fallas en sistemas de cómputo de alta prestaciones se puede notar la importancia de la unidad de E/S en las arquitecturas paralelas como un punto a mejorar para lograr cubrir las exigencias de las aplicaciones que utilizan HPC. Una manera de llevar a cabo este trabajo es utilizar técnicas de simulación para evaluar el efecto de los cambios de los factores con mayores influencias en las prestaciones del sistema de E/S paralelo.

Se puede disminuir la complejidad y la probabilidad de errores en la generación de sistemas híbridos desarrollando una simulación específica de éstos utilizando diferentes frameworks [5] [6] [7].

Las aplicaciones científicas con un uso intensivo de datos utilizan software de E/S paralelo para acceder a archivos. Contar con una herramienta que permita predecir el comportamiento de este tipo de aplicaciones en HPC es de gran utilidad para los desarrolladores de aplicaciones paralelas. Por otro lado, ABMS ha sido utilizado para modelar problemas y sistemas complejos en diversas áreas de la ciencia.

Evaluar las prestaciones del subsistema de E/S con diferentes configuraciones y la misma aplicación

permite adaptar la configuración de E/S teniendo en cuenta el patrón de acceso de la aplicación. Pero también puede ser una gran ventaja analizar las necesidades de las aplicaciones antes de configurar el sistema físico. Una manera de predecir el comportamiento de las aplicaciones en el sistema de cómputo ante distintas configuraciones, es utilizando técnicas de modelado y simulación.

Se está desarrollando modelos e implementando una simulación de la arquitectura de E/S paralela, por medio de técnicas de simulación basadas en agentes o Sistemas Multi-Agente, (MAS-MultiAgentSystems), para evaluar el efecto de dimensionar el sistema de E/S o cambiar componentes como la red de almacenamiento, dispositivos de E/S, entre otros [8].

Simulación de arquitecturas de Cloud Computing

CloudSim [9] [10] es un Framework de simulación generalizado y extensible que permite el modelado, la simulación y la experimentación de diferentes infraestructuras y servicios de aplicaciones de Cloud Computing. Un ejemplo de utilización es la simulación de muchos centros de datos.

Su arquitectura consiste en entidades específicas que se representan como clases Java que pueden ser heredadas o instanciadas para simular experimentos. Estas clases representan centros de datos, hosts físicos, máquinas virtuales, servicios a ejecutar en los centros de datos, servicios en la nube de usuarios, redes internas centro de datos y consumo de energía de los hosts físicos y elementos de los centros de datos [11]. Además, CloudSim soporta la inserción dinámica de los elementos de simulación y proporciona aplicaciones de paso de

mensajes y la topología de la red del centro de datos.

Una definición importante de CloudSim es la de entidad. Una entidad es una instancia de un componente, que es una clase o un conjunto de clases que representan un modelo CloudSim (datacenter, host). El motor de simulación es capaz de simular el tiempo de ejecución de las apps ingresadas como Cloudlets con información básica [12].

La versatilidad de CloudSim es la principal ventaja del sistema. La integración de nuevos parámetros y conceptos de la simulación es implementada desde abstracciones preestablecidas convenientemente por los autores. Las abstracciones principales son SimEvent [13], SimEntity [14], DataCenterCharacteristics y Vm.

En cada una de estas instancias es posible establecer el detalle del comportamiento y el estado de la simulación teniendo en cuenta el alcance del modelado que se planifica.

El aporte de un desarrollo de nuevos actores al componente Vm (Virtual Machine) posibilita que por medio de simulación se obtengan métricas de entrada/salida SAAS, PAAS o IAAS desarrollada con objetos que heredan de SimEvent y de SimEntity. En éste desarrollo se obtienen más detalles a la integración de éstos elementos en las características del objeto Vm que es el componente principal en los sistemas Cloud. Las estadísticas que aporta la nueva implementación de Vm, que se denomina SyntethicVm, son de gran interés ya que dan soporte a un espacio de memoria ram en tiempos de simulación. Dicho espacio de memoria principal está controlado por otro componente de SyntethicVm que es capaz de procesar instrucciones guardadas en el espacio de memoria sintético denominado RamEntity. Si las instrucciones se

guardan lógicamente, la nueva versión de Vm es capaz de administrar procesos en la nueva capa de ejecución [15].

Líneas de Investigación, Desarrollo e Innovación

Temas de Estudio e Investigación

- Arquitecturas multiprocesador para procesamiento paralelo: multiprocesador de memoria compartida, multiprocesador on-chip de memoria distribuida. Multicore, Clusters, Clusters de multicore. Grid. Cloud.
- Arquitectura de E/S paralela considerando el software, hardware, comunicaciones entre módulos y dispositivos de almacenamiento.
- Modelado y simulación basada en agentes.
- Nuevos aportes de desarrollos que mejoren los modelos de simulaciones con CloudSim para el análisis de la performance en sistemas de arquitecturas de software de Cloud Computing.

Resultados y Objetivos

Investigación experimental

- Diseño y desarrollo de modelos mediante técnicas de Modelado y simulación basada en agentes (ABMS) para analizar el comportamiento de las distintas capas de la pila de software de E/S.
- Utilización de agentes para generar la funcionalidad de los elementos físicos (procesadores, memoria, buses, drivers, entre otros) como así también

de las interfaces en las arquitecturas de E/S.

- Análisis y modelado de librerías de archivos para aplicaciones que utilizan cómputo de altas prestaciones-HPC. Se ha utilizado Amazon Web Services para creación de cluster virtuales y obtener métricas de la pila de software de E/S.
- Desarrollo de pruebas de conceptos con el modelo inicial del sistema de archivos paralelos utilizando el framework NetLogo.
- Obtención de un método de desarrollo de nuevos actores genéricos CloudSim que mejoran el modelado y la producción de estadísticas virtuales.
- Implementación de soportes de memoria principal en tiempos de simulación. Se los activa como componentes de las máquinas virtuales procesadoras de cloudlets.
- Implementación de la entidad que procesa instrucciones en el espacio de memoria de las máquinas virtuales. Se vinculan exitosamente los tiempos de procesamiento de un cloudlet y las instrucciones asignadas a las máquinas virtuales. El objetivo a mediano plazo es el contraste con las estadísticas de un despliegue de hardware que proporcione estadísticas de un sistema cloud computing privado.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D se participa en el dictado de la carrera de Ingeniería en Informática de la UNAJ.

También aportan trabajos de alumnos de las materias Redes de Computadoras 2 y Programación en Tiempo Real.

Durante 2017 se han realizado 2 publicaciones nacionales y 1 internacional.

En esta línea de I/D existe cooperación a nivel nacional e internacional. Hay un investigador realizando su Doctorado, 2 becarios EVC CIN y 3 alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing". Second Edition. Pearson Addison Wesley, 2003.
2. H Hennessy, J. L., Patterson, and D. A., Computer Architecture, Fourth Edition: A Quantitative Approach. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2006.
3. J. M. May, Parallel I/O for high performance computing. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2001.
4. V. Balaji, Earth system modelling – Volume 4. IO and Postprocessing. Springer, 2013.
5. D. Encinas, Utilización de un reloj global para el modelado de un ambiente simulado distribuido. XVIII Congreso Argentino de Ciencias de la Computación. 2012
6. D. Encinas, Simulación de una red CAN para dimensionar las comunicaciones de una IMU. VII Congreso Argentino de Tecnología Espacial. 2013.
7. D. Black, SystemC: From the Ground Up. Second Edition, Springer, 2010.
8. D. Encinas et al., Modeling I/O System in HPC: An ABMS Approach. The Seventh International Conference on Advances in System Simulation (SIMUL), ISBN: 978-1-61208-442-8, 2015.
9. R. Calheiros, R. Ranjan, A. Beloglazov, C. De Rose and R. Buyya "CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms" Published online 24 August 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/spe.995.
10. <http://www.cloudbus.org/cloudsim> 2018.
11. Hamza Ouarnoughi, Jalil Boukhobza, Frank Singhoff, Stephane Rubini, Erwann Kassis. "Considering I/O Processing in CloudSim for Performance and Energy Evaluation". OpenStack Cloud Software: Open source software for building private and public clouds. © Springer International Publishing AG 2016 M. Tauber et al. (Eds.): ISC High Performance Workshops 2016, LNCS 9945, pp. 591–603, 2016. DOI: 10.1007/978-3-319-46079-6_40.
12. Kushang Parikh, Nagesh Hawanna, Haleema. P.K, Jayasubalakshmi.R and N.Ch.S.N.Iyengar. School of Computing Science and engineering Vellore Institute of Technology, Tamil Nadu, "Virtual Machine Allocation Policy in Cloud Computing Using CloudSim in Java." 2015.
13. www.icsa.inf.ed.ac.uk/research/groups/hase/simjava 2018.
14. F. Howell, R Mc Nab. A discrete event simulation library for java. International Conference on Web-Based Modeling and Simulation. 1998.
15. D. Rosatto, R. Bond, M. Belizán, M. Morales, D. Encinas. Modelado y simulación de arquitecturas de Cloud Computing con CloudSim: comunicación entre entidades. XXIII Congreso Argentino de Ciencias de la Computación. ISBN 978-950-34-1539-9. 2017

Performance de cloud computing para HPC: despliegue y seguridad.

Brian Galarza¹, Gonzalo Zaccardi^{1,2}, Maximiliano Belizán¹, Jorge Osio¹, David Duarte¹,
Martin Morales^{1,3}, Diego Encinas^{1,4}

¹Instituto de Ingeniería y Agronomía - Universidad Nacional Arturo Jauretche

²Dirección Tecnológica – Universidad de Palermo

³Unidad CodApli - Facultad Regional La Plata - UTN

⁴Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática -
UNLP – Centro Asociado CIC

bgalarza@unaj.edu.ar, gzacca@palermo.edu, maximiliano.h.belizan@gmail.com,
josio@unaj.edu.ar, davito.duarte.22@gmail.com, martin.morales@unaj.edu.ar,
dencinas@unaj.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio de la performance de las arquitecturas tipo cloud a través del despliegue de IaaS y utilización de IaaS públicos, en particular en el área de cómputo paralelo de altas prestaciones (HPC). Enfocando a la obtención de herramientas que permitan predecir la eficiencia del sistema ante posibles escenarios. Analizando los diferentes componentes del sistema que pueden influir en las prestaciones significativamente y pueden llegar a modelarse y/o configurarse.

Palabras clave: *Arquitecturas Multiprocesador. Cloud Computing. OpenStack. Simulación.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación

“Modelado y Simulación en Cómputo de Altas Prestaciones (HPC). Aplicaciones en arquitecturas multiprocesador, sistemas paralelos y redes de datos” de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 186/15. Además, el Proyecto integra el Programa “Tecnologías de la información y la comunicación (TIC) en aplicaciones de interés social” de la UNAJ.

Por otra parte, se tiene financiamiento en el marco del programa “Universidad, Diseño y Desarrollo Productivo” del Ministerio de Educación a través de varios proyectos aprobados en la UNAJ.

Introducción

Cloud Computing es un paradigma que está en constante crecimiento durante estos últimos años, cada vez más compañías y grupos de investigación trabajan en conjunto con el fin de explotar las oportunidades ofrecidas por el mismo

[1]. Dicho paradigma ofrece muchas ventajas, tales como el bajo costo de implementación, ya que no se necesitan computadoras de última tecnología debido a que éstas trabajan conjuntamente (Clustering) con la posibilidad de escalar horizontalmente de manera sencilla. Además, hay software Open Source disponible para los nodos en el clúster como las infraestructuras Eucalyptus, OpenNebula, CloudStack u OpenStack integradas con GNU/Linux y compatibles, por ejemplo, con Amazon WebServices.

Despliegue de IaaS

Las comunicaciones en Cloud Computing son una parte fundamental del paradigma que consisten en utilizar distintos nodos y lograr hacerlos funcionar conjuntamente.

Para lograr una comunicación entre estos nodos se propone utilizar OpenStack [2].

OpenStack es un software de código abierto que permite la implementación de, por ejemplo, una “Infraestructure as a Service” (IaaS) a través de múltiples servicios que, de manera coordinada, cumplen diferentes propósitos para lograr el correcto funcionamiento de dicha infraestructura. Algunos de los servicios ofrecidos por OpenStack son: hypervisor (Nova), autenticación (Keystone), Imágenes (Glance), Dashboard (Horizon) y block storage (Cinder). Según las necesidades se pueden requerir de ciertos servicios u otros. La Arquitectura básicamente consiste en dos tipos de nodos: “Compute Node” y “Controller Node”. Se llaman Compute Node a todos aquellos que se encargan del procesamiento de servicios específicos mientras que Controller Node es aquel que comunica a cada uno de los anteriores [3] [4] [5].

La implementación de estas infraestructuras ofrece ventajas en las cuales los clústers virtualizados trabajan en conjunto ofreciendo un buen rendimiento a bajos costos y con posibilidad de escalabilidad al poder agregar mayor cantidad de nodos para procesamiento de manera sencilla.

Fuel es una herramienta Desarrollada por Mirantis en la cual se ejecuta un script que permite configurar, de manera más amigable respecto a OpenStack, los recursos que se desean otorgar a la infraestructura, como la cantidad de nodos, los núcleos de procesador, la memoria RAM, entre otros [6].

Fuel trabaja con un nodo master el cual es el encargado de controlar a los nodos slaves que contendrán la infraestructura OpenStack. Es decir, desde el nodo Fuel Master se indican qué paquetes se van a instalar en cada nodo slave (Glance, Nova-Compute, Keystone, etc.) para luego en los slaves tener armados los nodos compute y controller, sin necesidad de realizar configuraciones manuales en cada uno de los mismos.

Seguridad en arquitecturas de Cloud Computing

Aunque Cloud Computing tiene muchos beneficios, el uso inadecuado puede conllevar a pérdidas de datos sensibles para el usuario y el proveedor. Entre los principales riesgos y problemas de seguridad/privacidad se puede destacar los siguientes: Falta de control de datos, Ambigüedad de responsabilidad entre el usuario y el proveedor, autenticación y autorización, error de aislamiento, cumplimiento y riesgos legales, manejo de incidentes de seguridad, vulnerabilidad de la interfaz de administración, protección de aplicaciones y de datos, indisponibilidad del servicio, bloqueo del proveedor, eliminación de los datos

inseguros o incompletos, visibilidad y auditoría y seguridad [7] con la virtualización (ataques entre MVs, MVs inactivas desactualizadas, falta de monitoreo y registro de MVs, Host con MVs seguras e inseguras y ataques DoS por uso no autorizado de recursos compartidos) [8].

Es importante crear un modelo de seguridad de datos, para prevenir los problemas antes mencionados, que proporcionen seguridad en los: datos para procesar (encriptación homomórfica, ejemplo: Unpadded RSA), datos de transmisión (certificados SSL, ejemplo: HTTPS), datos de almacenamiento (encriptación, ejemplo: AES), administración de claves (ejemplo: HSM), Autenticación y autorización (autenticación de dos factores, ejemplo: Token), protección de usuario (alentar el encriptado de datos sensibles antes de subir) y controles de uso (Firewalls) [9]. OpenStack cuenta con un servicio llamado Keystone que ayuda en la autenticación y autorización distribuida de clientes, esto con apoyo de la encriptación de la base de datos, comunicación sobre capas seguras y el uso de Firewall, como Iptables, puede ayudar a proveer un Cloud más confiable [10].

Finalmente se propone la implementación de seguridad en la capa física. Ésta, va más allá del uso de OpenStack, ya que se busca cierta restricción al acceso físico de los equipos proveedores del servicio Cloud de OpenStack.

Estas implementaciones, más las actualizaciones en los distintos dispositivos (equipos, router, etc.) y software (OpenStack, iptables, etc) que se utilicen.

Líneas de Investigación, Desarrollo e Innovación

Temas de Estudio e Investigación

- Arquitecturas multiprocesador para procesamiento paralelo: multiprocesador de memoria compartida, multiprocesador on-chip de memoria distribuida. Multicore, Clusters, Clusters de multicore. Grid. Cloud.
- Plataformas de software para implementar y administrar Clouds públicos, privados e híbridos.
- Seguridad y privacidad en Cloud Computing.

Resultados y Objetivos

Investigación experimental

- Implementación de un IaaS encargado de realizar operaciones en procesamiento paralelo aumentando la eficiencia y reduciendo los costes generados.
- Implementación de OpenStack Dashboard y de un sistema desarrollado para poder controlar/administrar de manera visual (web) y más básica cada uno de los servicios.
- Utilización de Fuel para administrar OpenStack como sistema de administración de nube (Cloud Computing) a partir de la infraestructura de 2 nodos compute y 1 controller [11].

- Lograr escalabilidad agregando nuevos nodos compute a la infraestructura obteniendo un mayor performance en el sistema.
- Analisis del rendimiento de un Cloud privado en la ejecución de instancias personalizadas.
- Ejecutar sobre el modelo de nodos implementado en OpenStack la distribución GNU/Linux Hetnux, desarrollada en la UNAJ en el marco del programa “Universidad, Diseño y Desarrollo Productivo 2014”.
- Ejecutar diferentes benchmarks en la infraestructura desplegada sobre OpenStack para así comparar los resultados obtenidos con infraestructuras similares montadas en Azure y Amazon y poder realizar un análisis del rendimiento en cada caso.
- Implementación de un modelo de seguridad en un IaaS.
- Comprobar las mejoras prácticas para la seguridad de Cloud y modificar el modelo de seguridad, de ser necesario, para una mayor robustez.
- Como resultado del modelo de seguridad se espera un servicio de infraestructura Cloud que proteja los servicios usados y los datos generados por los usuarios.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D se participa en el dictado de la carrera de Ingeniería Informática de la UNAJ. También aportan trabajos de alumnos de las materias Sistemas Operativos 1, Redes de Computadoras 2, Programación en

Tiempo Real y Organización y Arquitecturas de Computadoras.

Durante 2017 se han realizado 3 publicaciones nacionales.

En esta línea de I/D existe cooperación a nivel nacional. Hay 2 investigadores realizando carreras de postgrado, 2 becarios de grado (EVC CIN y entrenamiento CIC) y 2 alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Kondo, D., Javadi, B., Malecot, P., Cappello, F., Anderson, D. P.: “Cost-benefit analysis of Cloud Computing versus desktop grids”. In: IPDPS '09 Proceedings. IEEE International Symposium on Parallel and Distributed Processing, Washington, USA (2009).
2. OpenStack Cloud Software: Open source software for building private and public clouds. <http://www.openstack.org>. Febrero 2015.
3. Galarza, B.; Tuamá, C.; Zaccardi, G.; Encinas, D.; Morales, M. “Implementaciones de Cloud Computing y aplicaciones en el ámbito universitario”. I Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2013). Ciudad de Córdoba, Argentina.
4. Zaccardi, G.; Galarza, B.; Encinas, D.; Morales, M. “Implementación de Cloud Computing utilizando OpenStack”. II Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2014). Ciudad de San Luis, Argentina.
5. Galarza, B.; Zaccardi, G.; Encinas, D.; Morales, M. “Análisis de despliegue de una IaaS utilizando Openstack”. XXI Congreso Argentino de Ciencias de la Computación (CACIC 2015). Ciudad de Junín, Argentina.
6. OpenStack Deployment Fuel. <https://www.mirantis.com/products/miranti>

- s-openstack-software/openstack-deployment-fuel/. Febrero 2016
7. Cloud Standards Consumer Council <http://www.cloud-council.org/deliverables/CSCC-Security-for-Cloud-Computing-10-Steps-to-Ensure-Success.pdf>. Marzo 2017
 8. Mishra, A.; Mathur, R.; Jain, S.; Rathore, J. “Cloud Computing Security”. *International Journal on Recent and Innovation Trends in Computing and Communication* ISSN 2321–8169 Volume: 1 Issue: 1 36 – 39. 2013
 9. Mohamed, E.; Abdelkader, H.; El-Etriby, S. “Data Security Model for Cloud Computing”. *Journal of Communication and Compute*. ISSN: 1047-1062. 2013
 10. Balu, V.; Mary, L. “A Model of Security Architecture on Private Cloud Using OpenStack”. *International Journal on Recent and Innovation Trends in Computing and Communication* ISSN: 2321-8169 Volume: 3 Issue: 2 587-590. 2015.
 11. G. Zaccardi, B. Galarza, M. Morales, D. Encinas. Despliegue y ejecución de un cloud privado. IV Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2016). Ciudad de Salta, Argentina. ISSN 2347-0372. 2016

Sintonización automática de aplicaciones paralelo/distribuidas basadas en algoritmos evolutivos y evolución diferencial

Caymes-Scutari Paola^{1,2}, Bianchini Germán¹, Tardivo María Laura^{1,2,3},
Méndez-Garabetti Miguel^{1,2}

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido
Departamento de Ingeniería en Sistemas de Información
Facultad Regional Mendoza/Universidad Tecnológica Nacional
Rodríguez 273 (M5502AJE) Mendoza, +54 261 5244579

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

³Departamento de Computación, Universidad Nacional de Río Cuarto, Río Cuarto,
Córdoba, Argentina

pcaymesscutari@frm.utn.edu.ar, gbianchini@frm.utn.edu.ar, mmendez@mendoza-conicet.gob.ar,
lauratardivo@dc.exa.unrc.edu.ar

RESUMEN

El procesamiento paralelo se ha convertido en un recurso muy utilizado para el procesamiento de datos y la resolución de problemas en diversos campos de la ciencia y la ingeniería. Sin embargo, el paradigma paralelo conlleva una complejidad considerable dado que involucra un gran número de aspectos no funcionales (como granularidad, técnica de descomposición y/o asignación, grado de concurrencia, etc.) que pueden impactar negativamente en el rendimiento de las aplicaciones. Mitigar dicho efecto negativo requiere de un cierto dominio y experiencia en su manejo, lo cual suele resultar prohibitivo para el usuario no experto en informática y en paralelismo en particular. Es por ello que tanto el proceso de desarrollo de aplicaciones paralelas como el proceso de sintonización de las mismas constituyen importantes desafíos a la hora de abordar un problema de manera paralela y eficiente. El Proceso de Sintonización

comporta una serie de etapas sucesivas mediante las cuales las aplicaciones son instrumentadas, monitorizadas, analizadas y sintonizadas (o ajustadas) de acuerdo a las características del problema que presentan o al entorno de ejecución, con el fin de mejorar su rendimiento y hacer un mejor aprovechamiento de los recursos. Sin lugar a dudas, el proceso de sintonización requiere un grado de conocimiento de la aplicación, el paradigma paralelo y los problemas de sintonización que usualmente lo vuelve una tarea muy difícil y restrictiva para lograr un rendimiento adecuado, sobre todo cuando de usuarios no expertos se trata. El presente proyecto propone abordar la sintonización automática de aplicaciones paralelas basadas en Evolución Diferencial.

Palabras clave: Evolución Diferencial, Algoritmos Evolutivos, Sintonización Automática, HPC, Problemas de Rendimiento, Problem Solver.

CONTEXTO

Este proyecto se enmarca en una línea de investigación que se viene desarrollando en los últimos años en el ámbito del LICPaD (Laboratorio de Investigación en Cómputo Paralelo/Distribuido) en la UTN-FRM. La misma se centra en el desarrollo de un entorno denominado EDDSAP (Entorno de Especificación, Desarrollo y Sintonización de Aplicaciones Paralelas [2, 3]), que integra la automatización de los procesos de desarrollo y sintonización de aplicaciones paralelas basadas en diferentes *problem solvers*. En este proyecto particular, se propone trabajar sobre el problem solver Evolución Diferencial, y actualmente se espera obtener financiamiento a través del proyecto SIUTNME0004819 (UTN) el cual se encuentra en evaluación.

1. INTRODUCCIÓN

La Evolución Diferencial (ED) [7] constituye una metaheurística poblacional [4] muy utilizada para la resolución de problemas de optimización, basada en diferencia de vectores para el tratamiento de la población. Cada paso evolutivo incluye tres fases, a saber: la mutación, el cruzamiento y la selección/reemplazo. El operador de mutación es el que efectivamente aplica la diferencia de vectores entre los miembros de la población corriente para determinar el grado y la dirección de los individuos resultantes. Se han estudiado y analizado las diferentes posibilidades de paralelización de ED, tanto para potenciar la calidad de los resultados obtenidos como la velocidad de procesamiento del método, lo cual ha incluido la revisión del estado del arte en la materia. En consecuencia, hemos considerado tres posibilidades de paralelización, basadas

en diferentes variantes del modelo de islas con migración [8], e implementadas sobre un modelo Master/Worker con topología de anillo (a uno o dos niveles). Una de las posibilidades se denomina **modelo subpoblacional**, pues una única población es dividida por el master en subpoblaciones procesadas por cada isla o worker. Otra de las posibilidades, denominada **modelo de islas clásico**, en la que cada isla maneja su propia población completa. La tercera posibilidad se denomina **modelo jerárquico de islas**, y compone los dos modelos anteriores de forma jerárquica. El algoritmo maneja múltiples poblaciones, que son procesadas por islas independientes, dotadas cada una de varios workers que realizan los cálculos. Los resultados obtenidos [9] permitieron concluir que el modelo subpoblacional reduce el tiempo total de ejecución, en detrimento de la calidad de las soluciones, que disminuye dado el tamaño de las subpoblaciones; que usando el modelo de islas clásico los resultados mejoran notablemente en detrimento del tiempo de ejecución; y que en el modelo jerárquico de islas el algoritmo alcanza una relación de compromiso entre la calidad de las soluciones y el tiempo de ejecución. Los tres modelos han sido implementados, y asimismo, se ha trabajado en la calibración a diferentes niveles de todos ellos a fin de ajustar su comportamiento y mejorar la calidad general del funcionamiento. Para ello fue diseñado un conjunto de experimentos a fin de ejecutar diferentes combinaciones de parámetros y problemas para cada prototipo desarrollado. Los resultados obtenidos han sido analizados a fin de establecer la incidencia de cada parámetro, lo cual resulta de suma importancia para la determinación de los parámetros factibles de ser sintonizados estática y/o dinámicamente y así

identificar los problemas de rendimiento típicos de las aplicaciones paralelas, y específicos de la ED. Uno de los objetivos más importantes dentro de este proyecto consiste en encapsular todo este conocimiento, mediante el modelado del algoritmo ED y la definición de un modelo de rendimiento del mismo, que provea los elementos necesarios para abordar la sintonización del método de forma automática y dinámica. Que la sintonización sea automática, significa que las técnicas, capacidades, y conocimiento de rendimiento estarán incorporadas en la propia aplicación de forma automática, y como consecuencia el usuario no tendrá necesidad de responsabilizarse de la sintonización, sino que resultará una capacidad incluida en la aplicación de forma transparente para el usuario, lista para ser utilizada. Ello representa una gran ventaja dado que le da acceso al usuario a la reutilización de conocimiento experto en el análisis y resolución de problemas de rendimiento, sin necesidad de involucrarse en ninguna etapa del complejo proceso de sintonización.

Previo el avance en tal sentido se ha seleccionado un caso de aplicación particular de ED: la predicción de incendios forestales, trabajando sobre la representación de la información y el procesamiento de la misma, y diseñando una serie de experimentos que permitan analizar el impacto de los diferentes parámetros en la calidad de la ejecución del caso de aplicación. Por otro lado, se realizó un estudio comparativo [6] entre la predicción de incendios utilizando esta versión paralela de ED frente a otros métodos ya conocidos de predicción de incendios forestales [1, 5], para así evaluar la efectividad y eficiencia del método, cuestión en la que se continúa trabajando, dado que dicha comparación

requiere además afrontar la sintonización del método para mejorar su efectividad.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Esta línea de investigación tiene como objetivo general el desarrollo de un Entorno de Especificación, Desarrollo y Sintonización de Aplicaciones Paralelas (EEDSAP). El mismo ha sido concebido como un entorno integral para el desarrollo y la sintonización de aplicaciones paralelas, que hace transparente los aspectos relacionados con el diseño de aplicaciones paralelas y el proceso de sintonización automática y dinámica de las mismas, por medio de la utilización de esqueletos, plantillas e instrumentación de rendimiento. Estas características lo hacen especialmente útil para usuarios no expertos en paralelismo, sintonización, y hasta en programación, dado que el entorno ofrece una interfaz declarativa, por lo que no se requieren conocimientos específicos de programación por parte del usuario. En términos de intervención del usuario en el ciclo de vida de la aplicación, EEDSAP propone una primera fase cooperativa en la que el usuario describe el problema, una segunda fase automática en la que se genera el código de la aplicación y se lo dota de capacidades para su sintonización a partir de la interpretación de un esqueleto con la información provista por el usuario en la especificación; finalmente, en una tercera fase dinámica (es decir durante la ejecución) los parámetros de ejecución de la aplicación son sintonizados o ajustados a fin de alcanzar un rendimiento superior y una utilización más eficiente de los recursos computacionales involucrados, de acuerdo al estado del entorno de

ejecución. En este proyecto se dará tratamiento al *problem solver* Evolución Diferencial a fin de extender esta aproximación y así ampliar el uso y utilidad de EEDSAP.

3. RESULTADOS ESPERADOS

El principal aporte que se espera alcanzar es la dotación del método evolución diferencial con la capacidad de ajustar automática y dinámicamente aquellos parámetros de funcionamiento que permitan alcanzar una ejecución más eficiente. Tal capacidad será heredada por todas aquellas aplicaciones que se generen automáticamente a partir de él, dado que tanto el método ED como su modelo de rendimiento serán encapsulados en un módulo de software que permita generar aplicaciones paralelas sintonizables basadas en ED, haciendo transparentes tanto los conceptos de paralelismo como los de sintonización a la hora de desarrollar la aplicación. En otras palabras, el aporte general que se espera consiste en ofrecer una herramienta que requiera del usuario mínimos conocimientos relacionados con el paradigma paralelo, lenguajes de programación, lenguajes de especificación, mientras que ofrezca un entorno amigable e integral tanto para el desarrollo como para la sintonización automática de aplicaciones paralelas basadas en ED. Ello constituirá una contribución importante también para la línea de investigación que enmarca a este proyecto, cuyo objetivo se centra en el desarrollo de un entorno que integre el desarrollo automático con la sintonización, donde se espera que ambos procesos resulten transparentes al usuario, a la hora de crear y utilizar aplicaciones paralelas, y como parte del cual se incorporará ED como un nuevo *problem*

solver de desarrollo y sintonización. Tales características resultan primordiales en una herramienta pensada para usuarios no expertos, tanto en el campo del paralelismo, como en el de la sintonización, e incluso de los propios métodos de resolución (sean algoritmos genéticos, evolución diferencial, o cualquier otro que se integre a EEDSAP). Otro aporte importante que se espera alcanzar es la definición de un modelo de rendimiento de ED que encapsule la representación de las características del funcionamiento del *problem solver*, y permita analizar los problemas de rendimiento que manifiesta.

4. FORMACIÓN DE RECURSOS HUMANOS

La temática propuesta por este proyecto permite continuar con la formación de los distintos integrantes del grupo de trabajo, de forma complementaria a la formación adquirida hasta el momento, ahora en el campo del análisis y sintonización de software basado en evolución diferencial. En el caso particular de la Lic. Tardivo y el Ing. Méndez, su participación en el proyecto, el estudio, experimentación y puesta en marcha de los procesos de análisis y sintonización de metaheurísticas como la evolución diferencial, será fundamental como elemento para su formación en el área de sintonización de aplicaciones paralelas, aplicable a sus respectivos planes de trabajo. Ambos estudiantes cursan el Doctorado en Ciencias de la Computación de la Universidad Nacional de San Luis, y revisten como becarios doctorales de CONICET. Asimismo, el grupo de trabajo siempre está abierto a la incorporación de nuevos integrantes (de grado o postgrado) que deseen

familiarizarse con las temáticas que aquí se describen.

5. BIBLIOGRAFÍA

- [1] Bianchini, G., Caymes Scutari, P., Méndez Garabetti, M., Evolutionary-Statistical System: a Parallel Method for Improving Forest Fire Spread Prediction, *Journal of Computational Science (JOCS, Elsevier)*. Vol. 6, pp. 58-66, 2015.
DOI: 10.1016/j.jocs.2014.12.001
- [2] P. Caymes Scutari, G. Bianchini, Environment for the Automatic Development and Tuning of Parallel Genetic Algorithms. *Proceedings of 40 JAIIO at the High-Performance Computing Symposium (HPC 2011, ISSN: 1851-9326) (2011) 17 a 20.*
- [3] P. Caymes Scutari, G. Bianchini, A. Sikora, T. Margalef, Environment for Automatic Development and Tuning of Parallel Applications. *Proceedings of the 2016 International Conference on High Performance Computing & Simulation (HPCS 2016) ISBN: 978-1-5090-2088-1/16/\$31.00 ©2016 IEEE pp. 743-750. DOI: 10.1109/HPCSim.2016.7568409.*
- [4] Goldberg, D.E., Genetic and evolutionary algorithms, *Come of age*, *Communications of the ACM*, 37 (3) pp. 113119, 1994.
- [5] Méndez-Garabetti, M., Bianchini, G., Tardivo, M. L., & Caymes-Scutari, P., Comparative analysis of performance and quality of prediction between ESS and ESS-IM. *Electronic Notes in Theoretical Computer Science*, 314, 45–60, 2015. DOI: 10.1016/j.entcs.2015.05.004
- [6] Tardivo, M.L., Caymes Scutari, P., Bianchini, G., Méndez Garabetti, M., Cencerrado, A., Cortés, A., A comparative study of evolutionary statistical methods for uncertainty reduction in forest fire propagation prediction. *Procedia Computer Science*, 108, 2018–2027, 2017.
DOI: 10.1016/j.procs.2017.05.252
- [7] R. Storn, K. Price, *Differential Evolution - A simple and efficient adaptive scheme for global optimization over continuous spaces*, Technical Report TR-95-012, Berkeley, CA (1995).
- [8] E.G. Talbi, *Metaheuristics: From Design to Implementation*, John Wiley & Sons, Hoboken, New Jersey (2009).
- [9] M.L. Tardivo, P. Caymes-Scutari, M. Méndez-Garabetti, G. Bianchini, Parameters Calibration for Parallel Differential Evolution based on Islands. *XIII Workshop Procesamiento Distribuido y Paralelo, CACIC 2013*. 296-305, 2013.

Garantizando la consistencia de JavaScript en un contexto de memoria compartida

Matías Teragni, Gonzalo Zabala, Ricardo Morán, Sebastián Blanco.

Centro de Altos Estudios en Tecnología Informática

Facultad de Tecnología Informática

Universidad Abierta Interamericana

Av. Montes de Oca 745, Ciudad Autónoma de Buenos Aires, República Argentina

(+54 11) 4301-5323; 4301-5240; 4301-5248

{ Matias.Teragni, Gonzalo.Zabala, Ricardo.Moran, Sebastian.Blanco}@uai.edu.ar

Resumen

El propósito de este proyecto es diseñar e implementar un conjunto de librerías, técnicas y funcionalidades que permitan, dado un sistema distribuido de memoria compartida, generar código javascript que ejecute haciendo uso de los diversos nodos presentes prestando garantías sobre la consistencia del programa ejecutado, y brindando las herramientas necesarias para poder garantizar la semántica del código escrito por el programador, y su predictibilidad a la hora de ejecutar.

Palabras clave: Programación

Distribuida, Consistencia, Javascript

Contexto

El presente proyecto será radicado en el Centro de Altos Estudios en Tecnología Informática, dependiente de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana. El mismo se encuentra inserto en la línea de investigación “Algoritmos y software”. El financiamiento está compartido entre el CONICET y la misma Universidad por partes iguales.

Introducción

Se entiende a un sistema distribuido como un conjunto de computadoras independientes, cada una con su propia memoria y capacidad de procesamiento, que se encuentran interconectadas por medio de una red, y sobre las cuales opera un conjunto de software que colabora para lograr un comportamiento complejo como un todo.

Los sistemas distribuidos pueden proveer, dada su naturaleza, un conjunto de beneficios, desde resistencia a fallos mediante replicación, hasta la capacidad de ejecutar procesos sumamente complejos en computadoras económicas mediante la partición del problema en partes más simples.

A la hora de construir sistemas distribuidos existen múltiples paradigmas que se pueden aplicar en función de las necesidades particulares del contexto. Estos varían en nivel de abstracción, y las interfaces y lenguajes que le permiten utilizar a los desarrolladores para construirlos.

Particularmente uno de los mecanismos

que tuvo un resurgimiento últimamente, gracias a los avances tecnológicos que compensan algunas de sus falencias, fue la idea de una memoria compartida entre los nodos.

La idea no es novedosa, pero las limitaciones en velocidades de red y capacidad de memoria de los distintos nodos volvieron a este modelo poco práctico en comparación a algunas de las alternativas.

Este modelo es muy utilizado en los servicios provistos en la “nube” ya que para garantizar tiempos de acceso, fault tolerance, y service level assurance, los servidores normalmente se replican en varios datacenters a lo largo del mundo.

Esta replicación trae consigo claros beneficios pero no está libre de riesgo, ya que la consistencia de la información presente en dichos nodos puede generar comportamientos indebidos en las aplicaciones de los clientes.

Estos problemas son de interés actual, y muchos de los esfuerzos se centran en tratarlos exclusivamente a nivel de datos, sin considerar las particularidades que le pueden generar estas inconsistencias a un programa que ejecute en función a esta información que puede ser inconsistente.

El presente trabajo tiene como objetivo desarrollar un conjunto de librerías que permitan mantener las ventajas de las arquitecturas distribuidas pudiendo escribir código cuya ejecución sea predecible y garantizable.

Líneas de Investigación, Desarrollo e Innovación

Para poder abordar este problema existen dos conflictos en particular que deben tratarse. Por un lado está el problema de la replicación de los datos entre los datos (memoria) entre los distintos nodos, y por otro la construcción de herramientas que permitan garantizar la semántica del código escrito por un programador.

En cuanto a la replicación respecta, es un problema que ya fue abordado tanto por la comunidad como por los autores de este artículo con diversos grados de éxito. Con respecto a garantizar la semántica de un código ejecutado de manera distribuida existen algunas restricciones que se deben poder garantizar sobre este proceso.

- **Consistencia Eventual:** Se refiere a la capacidad que tiene un sistema distribuido de que todos los nodos que lo integran tengan una convergencia sobre el estado de la información compartida. El nombre hace referencia a que dicha convergencia puede no ser alcanzada de forma instantánea, pero dado tiempo infinito es demostrable que la misma existe.
- **Consistencia Causal:** Se refiere a la capacidad que tiene un sistema distribuido de preservar órdenes parciales de lecturas/modificaciones ante dependencia funcional de las acciones realizadas. Esto es fundamental para garantizar la semántica de cualquier programa que deba ejecutarse.

Dado un sistema de replicación que pueda garantizar ambas características consideramos que puede construirse sobre

el mismo un conjunto mínimo de operaciones que permitan a un programador tener un control fino de su modelo de ejecución, permitiéndole generar secciones críticas, mutex, y semáforos cuando sea necesario, y aprovechar las características concurrentes inherente a un sistema distribuido cuando no lo sea. Preservando así la simplicidad del código, obteniendo algunos de los beneficios de operar en un espacio distribuido pero sin sacrificar la semántica del código, evitando la generación de race conditions, ejecuciones secuencialmente inválidas, ni comportamientos impredecibles en un código que no lo amerite.

Resultados y objetivos

El objetivo de este proyecto es diseñar e implementar un esquema control que permita a un programador definir inequívocamente las necesidades de sincronismo que tiene el código que genera.

Para los prototipos se propone utilizar javascript dada su gran popularidad, y su naturaleza multiplataforma. Un requisito adicional es no imponer en el motor de ejecución del lenguaje ninguna modificación, permitiendo que estos beneficios sean aplicables as-is en cualquier contexto donde código javascript pueda ejecutar.

Una vez demostrados formalmente estos mecanismos, y planteados prototipos de su implementación el siguiente paso será evaluar extender estas funcionalidades a otros lenguajes de programación que implementan más restricciones, como podría ser Java, C# o C.

Formación de Recursos Humanos

El equipo de trabajo está conformado por un investigador adjunto del Centro de Altos Estudios en Tecnología Informática (CAETI) quien ejerce el rol de director del proyecto, dos doctorandos, y un ayudante alumno de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana.

Referencias

- LeBlanc, T.J; Markatos, E.P. (1992). Shared memory vs. message passing in shared-memory multiprocessors. Proceedings of the Fourth IEEE Symposium on Parallel and Distributed Processing
- Coulouris, George; Jean Dollimore; Tim Kindberg; Gordon Blair (2011). Distributed Systems: Concepts and Design (5th Edition). Boston: Addison-Wesley.
- Ngo, T.A; Snyder, L. (1992). On the influence of programming models on shared memory computer performance. Scalable High Performance Computing Conference, 1992. SHPCC-92, Proceedings.
- LeLann, G. (1977). "Distributed systems - toward a formal approach,". Information Processing. 77: 155-160. – via Elsevier.
- Andrews, Gregory R. (2000), Foundations of Multithreaded, Parallel, and Distributed Programming, Addison-Wesley, ISBN 0-201-35752-6.
- Lamport, L. (1977). Proving the Correctness of Multiprocess Programs. IEEE Transactions on Software Engineering (Volume: SE-3, Issue: 2, March 1977)

Simulación computacional, ciencia de los datos, cómputo de alto rendimiento y optimización aplicados a mejorar la predicción de modelos de simulación que representan la evolución de sistemas complejos

Mariano Trigila², Adriana Gaudiani^{1,4}, Emilio Luque³, Marcelo Naiouf¹

¹ Instituto de Investigación en Informática LIDI, Facultad de Informática, Universidad Nacional de La Plata - Centro Asociado de la CIC
mnaiouf@lidi.info.unlp.edu.ar

² Facultad de Ingeniería y Ciencias Agrarias, Pontificia Universidad Católica Argentina
marianotrigila@gmail.com

³ Depto. de Arquitectura de Computadores y Sistemas Operativos, Universidad Autónoma de Barcelona, 08193 Bellaterra (Barcelona) España
emilio.luque@uab.es

⁴ Instituto de Ciencias, Universidad Nacional de General Sarmiento.
agaudi@ungs.edu.ar

Resumen

El objetivo principal de esta propuesta es mejorar de manera automática la calidad de la simulación de un sistema dinámico complejo, proponiendo una metodología computacional de calibración.

En particular se utilizará con un modelo de translación de ondas a lo largo del cauce de un río [9]. Esta metodología se centrará en manejar la incertidumbre en los parámetros del modelo para optimizar la calidad de la predicción del simulador.

Este trabajo continúa con la línea de investigación previa donde se propuso una metodología computacional de optimización que fue validada con este simulador [6]. En esa oportunidad se elaboró una mejora computacional independiente del sistema simulado, pero en este trabajo se aprovecharán las características propias del dominio del sistema para calibrar el modelo. Se aprovecharán las propiedades de localidad en los valores de los parámetros, proponiendo una calibración en pasos sucesivos, en tramos seleccionados sobre el cauce del río.

La implementación de este trabajo requiere la utilización de técnicas de las ciencias de los datos, del cómputo de alto rendimiento y métodos del campo de la optimización [7] [10], buscando lograr el mayor ahorro en el uso de recursos computacionales.

Palabras Claves

Optimización y simulación, calibración automática, ciencia de los datos, cómputo eficiente, cómputo de alto rendimiento.

Contexto

El desarrollo de este trabajo se construye en base a la relación de colaboración establecida con los siguientes equipos de investigación.

- III-LIDI, Facultad de Informática, Universidad Nacional de La Plata.
- High Performance Computing for Efficient Applications and Simulation Research Group (HPC4EAS), Universidad Autónoma de Barcelona.
- Instituto de Ciencias de la Universidad Nacional de General Sarmiento.

- Proyecto LAPP: Análisis y exploración de visualización de datos espacio-temporales en el dominio de hidrología. Pontificia Universidad Católica Argentina.

- Instituto Nacional del Agua (INA), Laboratorio de Hidráulica Computacional (LHC).

- Group on Non-Equilibrium Processes and Transport Phenomena. Instituto de Ciencias de la Universidad Nacional de General Sarmiento.

El trabajo comenzó en 2017 y se lleva adelante en el ámbito de una tesis doctoral en Ciencias Informáticas, en la UNLP. El grupo que acompaña el proyecto viene trabajando en este tema desde 2010, y de estas investigaciones resultó una tesis doctoral en la que se basa esta propuesta, y varias publicaciones [5] [11]. Los resultados esperados son de interés para los grupos con los que se mantiene colaboración.

Introducción

Las inundaciones por desborde de ríos, cuyo impacto involucra a la sociedad, son eventos que afectan los recursos económicos en las zonas de influencia de subidas o bajantes del agua, incluso ponen en riesgo las vidas humanas. La simulación y predicción de estos eventos juegan un rol primordial como herramientas para brindar alertas. Estas son las causas principales de la elección de trabajar, en este proyecto, sobre mejorar la certeza provista en los datos simulados que proveen los modelos hidrológicos. En particular, los que se utilizan para predecir los niveles del agua del río Paraná, en Argentina, y en consecuencia, para predecir inundaciones en su zona de influencia [1] [9].

Las inundaciones de ríos de planicie es un fenómeno natural que ocurre con recurrencia, y con el cambio del clima

global se espera que acontecimientos extremos más severos y más frecuentes tengan ocurrencia en un futuro cercano [1] [8].

Por su propia naturaleza, el cauce y el lecho del río suelen modificarse con el transcurrir del tiempo y con ello se modifican los contornos y las áreas inundables, y en consecuencia se modifican los parámetros que caracterizan los modelos que lo representan, ya que se está en presencia de un sistema físico, complejo y dinámico. Este dinamismo le agrega mayor incertidumbre a la predicción que los especialistas realizan utilizando los datos simulados que se obtienen de los modelos computacionales de dicho sistema [4] [12] [13].

La Ciencia de la Computación y las tecnologías involucradas proveen los métodos y técnicas que brindarán los aportes necesarios para optimizar el proceso de simulación de estos sistemas [1]. En consecuencia, se podría ofrecer un aporte que mejore los pronósticos sobre el comportamiento de dichos sistemas físicos. Se propone aportar una mejora en la predicción de la simulación de un modelo computacional utilizado para brindar alertas a la población sobre crecidas y bajantes del cauce de un río, con gran impacto social y económico en la población costera [5] [10]. Se propone ofrecer una metodología de calibración automática del simulador, que se adapte a períodos de creciente o de bajante del río, debido a que estos eventos incrementan fuertemente el porcentaje de error en la predicción [8]. Se espera obtener amplios beneficios en las mejoras propuestas mediante los avances en la Ciencia de los Datos y en la Optimización vía Simulación, como también en el uso de técnicas más modernas del HPC [2] [3] [7]. En particular se optimizará el uso de recursos computacionales al implementar esta metodología.

Líneas de Investigación, Desarrollo e Innovación

Este trabajo está enmarcado por las líneas de investigación:

a) *computación eficiente y segura para la simulación*

b) *optimización de aplicaciones y optimización del uso de arquitecturas paralelas híbridas y cloud.*

Ambas son líneas de investigación relacionadas con los trabajos de los grupos de colaboradores nombrados en el contexto de esta presentación.

El desarrollo del mismo contempla ir cumpliendo con los siguientes objetivos:

- Estudiar el modelo de simulación del INA (Instituto Nacional del Agua)¹, utilizado para brindar alertas sobre el cauce del Río Paraná, en la Cuenca del Plata, mediante un estudio exhaustivo de la estructura de sus datos, de los datos de entrada y salida del modelo y del comportamiento del simulador [8][9].

- Estudiar el problema de optimización y calibración automática del simulador seleccionado para este trabajo, como también el problema de la búsqueda de patrones de comportamiento del río seleccionado en función de eventos extraordinarios, con el objetivo de analizar períodos de validez de las soluciones encontradas [6][12].

- Analizar y aplicar las posibles opciones que brinden una metodología de optimización vía simulación de la predicción, en modo dinámico, que se

adapte a la mayor extensión posible del dominio del río (incrementando la cantidad de parámetros de ajuste todo lo posible), con el objetivo de manejar más eficientemente los errores en la predicción que se originan por subidas o bajantes bruscas del nivel del cauce [2] [5] [13].

- Estudiar y aplicar las técnicas y metodologías que ofrece la programación en paralelo y el HPC en la implementación de la metodología propuesta.

- Analizar y utilizar las tecnologías, plataformas, frameworks y lenguajes de programación necesarios para diseñar e implementar las soluciones que sean necesarias para llevar adelante las experiencias requeridas durante el desarrollo de este trabajo.

El uso de los avances de HPC será imprescindible tanto para crear soluciones eficientes como para llevar adelante las experiencias de validación de la metodología propuesta. Se requiere una gran carga de cómputo al necesitar múltiples ejecuciones del simulador y al procesamiento de las enormes cantidades de datos que se obtendrán [1][3][7]. Se trabajará especialmente en minimizar los recursos computacionales, como tiempo de uso de CPU's, pensando en el mayor ahorro energético posible.

Resultados y Objetivos

Describir una metodología o una serie de procedimientos particulares basados en ciencias de los datos, desarrollo de heurísticas y simulación computacional, con aplicación de los métodos de la computación de alto rendimiento. Se espera proveer certeza en la capacidad de predicción del modelo, mediante un método de calibración automática; especialmente ante eventos extremos de

¹ G. Latessa, "Modelo hidrodinámico del río Paraná para pronóstico hidrológico: Evaluación del performance y una propuesta de redefinición geométrica.," INA - UBA, Buenos Aires , 2011

subida o bajante del cauce del río, siendo en estos momentos (de sequía o de lluvias) durante los que se desencadenan importantes desajustes al modelo de simulación.

Desarrollar herramientas de soporte necesarias para llevar adelante las investigaciones de este trabajo, con posibilidad de ofrecer herramientas creativas e innovadoras que puedan utilizarse en la toma de decisiones en alguna de las áreas de incumbencia de este proyecto.

Describir una serie de procedimientos generales basados en Ciencia de la Computación y Ciencias de los Datos para proveer un modelo de optimización de la predicción.

Estudiar la factibilidad de extender la metodología desarrollada para optimizar la predicción de otro modelo de simulación de fenómenos naturales.

Formación de Recursos Humanos

El desarrollo de este trabajo corresponde al ámbito de la tesis doctoral en Ciencias Informáticas de la UNLP, llevada adelante por Mariano Trigila bajo la dirección del Dr. E. Luque y la asesoría científica de la Dra. Gaudiani.

En el transcurso de este año se presentará la solicitud de becas de investigación otorgada por la UNGS a estudiantes avanzados de la carrera de Licenciatura en Sistemas para participar en temas de investigación relacionados con esta propuesta.

Referencias

[1] S. Balica, "Parametric and physically based modelling techniques for flood risk and vulnerability assessment: A comparison," *Environmental*

Modelling & Software, vol. 41, pp. 84-92, 2013.

[2] E. Cabrera, M. Taboada, M. Iglesias, F. Epelde y E. Luque, "Simulation Optimization for Healthcare Emergency Departments," *Procedia Computer Science*, vol. 9, pp. 1464-1473, 2012.

[3] A. Chi Zhou, B. He y I. Shadi, "A Taxonomy and Survey on eScience as a Service in the Cloud," *Cornell University Library*, 2014.

[4] A. Domeneghetti, A. Castellarin y A. Brath, "Assessing rating-curve uncertainty and its effects on hydraulic model calibration," *Hydrology and Earth System Sciences*, vol. 16, pp. 1191-1202, 2012.

[5] A. Gaudiani, E. Luque, P. García, M. Re, M. Naiouf y A. De Giusti, "How a Computational Method Can Help to Improve the Quality of River Flood Prediction by Simulation.," *Advances and New Trends in Environmental and Energy Informatics. Springer.*, pp. 337-351, 2016.

[6] A. Gaudiani, "Simulación y optimización como metodología para mejorar la calidad de la predicción en un entorno de simulación hidrográfica," *Tesis de doctorado - UNLP*, 2015.

[7] T. Hey, S. Tansley y K. Tolle, "The

- Fourth Paradigm: Data-Intensive Scientific Discovery,” *Proceedings of the IEEE*, vol. 99, n° 8, pp. 1324-1337, 2009.
- [8] A. Menéndez, “Three decades of development and application of numerical simulation tools at INA Hydraulics Lab,” *Mecánica Computacional*, vol. 21, pp. 2247-2266, 2002.
- [9] M. Re, N. Badano, E. Lecertúa, F. Re y A. Menéndez, “Modelización matemática de una cuenca de llanura extensa,” *Mecánica Computacional*, vol. 27, pp. 351-368, 2008.
- [10] J. F. Santucci y L. Capocchi, “Optimization via Simulation of Catchment Basin Management Using a Discrete-event Approach,” *Simulation*, vol. 91, pp. 43-58, 2015.
- [11] M. Trigila, “Cluster LAPP: Implementación de un cluster de bajo costo para procesamiento paralelo” de *World Congress on Systems Engineering and Information Technology - WCSEIT'2015*, Vigo, España, 2015.
- [12] J. Warmink, J. Janssen, M. Booij y M. Krol, “Identification and classification of uncertainties in the application of environmental models,” *Environmental Modelling & Software*, vol. 25, n° 12, pp. 1518-1527, 2010.
- [13] Q. Wu, S. Liu, Y. Cai, X. Li y Y. Jiang, “Improvement of hydrological model calibration by selecting multiple parameter ranges,” *Hydrology and Earth System Sciences.*, vol. 21, pp. 393-407, 2017.

ARQUITECTURAS MULTIPROCESADOR EN HPC: SOFTWARE, MÉTRICAS, MODELOS Y APLICACIONES.

De Giusti Armando^{1,2}, Tinetti Fernando^{1,3}, Naiouf Marcelo¹, Chichizola Franco¹, De Giusti Laura^{1,3}, Villagarcía Horacio^{1,3}, Montezanti Diego¹, Encinas Diego¹, Pousa Adrián¹, Rodríguez Ismael¹, Rodríguez Eguren Sebastián¹, Iglesias Luciano¹, Paniego Juan Manuel¹, Pi Puig Martín¹, Libutti Leandro, Dell'Oso Matías¹, Mendez Mariano^{1,4}

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 115, La Plata, Buenos Aires

Comisión de Investigaciones Científicas de la Pcia. de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CIC - Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

⁴Facultad de Ingeniería - Universidad de Buenos Aires

{degiusti, fernando, mnaiof, francoch, ldgiusti, hvw, dmontezanti, dencinas, apousa, ismael, seguren, li, mmendez, jmpaniego, mpipuig, llibutti, mdelloso}@lidi.info.unlp.edu.ar, marianomendez@gmail.com

Con la colaboración en la dirección de Tesis de Posgrado de la Universidad Autónoma de Barcelona (España) y la Universidad Complutense de Madrid (España).

RESUMEN

La línea presentada en este trabajo tiene como eje central las arquitecturas paralelas para Cómputo de Altas Prestaciones, con los siguientes objetivos:

- Caracterizar las arquitecturas multiprocesador orientadas a computación de alto desempeño, analizando técnicas para el desarrollo de código eficiente sobre las mismas y métricas de rendimiento computacional y energético. Principalmente teniendo en cuenta aquellas arquitecturas que son combinaciones de cluster, multicores, aceleradores (GPUs, FPGAs, Xheon Phi), placas de bajo costo (Raspberry PI, Odroid).

- Estudiar la conformación de clusters y clouds a partir de configuraciones homogéneas e híbridas de multiprocesadores. Y analizar la performance de aplicaciones sobre los mismos, considerando eficiencia computacional/energética y escalabilidad, así como la tolerancia a fallos.

- Analizar y desarrollar software de base para estas arquitecturas con el objetivo de optimizar del rendimiento y consumo energético en aplicaciones de propósito general.

- Estudiar clases de aplicaciones inteligentes en tiempo real, en particular el trabajo colaborativo de robots conectados a un cloud y procesamiento de Big Data.

Es de hacer notar que este proyecto se coordina con otros proyectos en curso en el III-LIDI, relacionados con Computación de Alto Desempeño, Algoritmos Paralelos, Sistemas Distribuidos y Sistemas de Tiempo Real.

Palabras claves: *Sistemas Paralelos – Multicore – GPU – FPGAs - Cluster y Cloud Computing - Cluster híbridos - Performance y eficiencia energética - Tolerancia a fallas – Scheduling - Cloud Robotics - Big Data.*

CONTEXTO

Esta línea de Investigación está dentro del proyecto: “Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real.” acreditado en el marco del Programa de Incentivos del Ministerio de Educación, y de proyectos específicos apoyados por organismos nacionales e internacionales. También del proyecto “Computación de Alto Desempeño, Minería de Datos y Aplicaciones de interés social en la Provincia de Buenos Aires” financiado por la CIC PBA dentro de la convocatoria a Proyectos de Innovación y Transferencia en Areas Prioritarias de la Pcia. de Buenos Aires (PIT-AP-BA). El III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD) del MINCYT y en esta línea de I/D hay cooperación con varias Universidades de Argentina, de América Latina y Europa en proyectos con financiación nacional e internacional.

En la Facultad de Informática de la UNLP (a partir del equipo del proyecto) se dictan asignaturas optativas en la currícula de grado de las carreras de Licenciatura en Informática, Licenciatura en Sistemas e Ingeniería en Computación relacionadas con Cloud Computing, Programación sobre GPGPUs y procesamiento de Big Data. Además, la Facultad aprobó y financia el proyecto “Transformación de Algoritmos para Nuevas Arquitecturas Multiprocesador”.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática, así como el desarrollo de la Maestría y Especialización en Computación de Altas Prestaciones, acreditadas por CONEAU.

Por último, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado y se ha tenido el apoyo de diferentes empresas (IBM, Microsoft, Telecom, INTEL, AMAZON AWS) en las temáticas de Cloud Computing y Big Data.

1. INTRODUCCIÓN

La investigación en Paralelismo (a partir de arquitecturas multiprocesador distribuidas o concentradas en supercomputadoras) es una de las líneas de mayor desarrollo en la Ciencia Informática actual [GRA03]. La utilización de clusters, multiclusters, grids y clouds, soportadas por redes de diferentes características y topologías se ha generalizado, tanto para el desarrollo de algoritmos paralelos orientados a HPC como para el manejo de aplicaciones distribuidas y/o servicios WEB concurrentes [GRA03][MCC12]. El área de cómputo de alto desempeño (o de altas prestaciones) se ha convertido en clave debido al creciente interés y necesidad por el desarrollo de soluciones a problemas con muy alta demanda computacional y de almacenamiento, convirtiéndose en ocasiones en la única alternativa real y produciendo transformaciones profundas en las líneas de I/D [RAU10][KIR12]. El rendimiento en este caso está relacionado con dos aspectos: por un lado, las arquitecturas de soporte y por el otro los algoritmos que hacen uso de las mismas, y el desafío se centra en cómo aprovechar las prestaciones obtenidas a partir de la evolución de las arquitecturas físicas.

Tradicionalmente el objetivo principal del cómputo de altas prestaciones (HPC, *High-Performance Computing*) fue mejorar el rendimiento y ocasionalmente la relación precio/rendimiento, donde rendimiento hace referencia a la velocidad de procesamiento. La constante búsqueda por mejorar la velocidad de procesamiento llevó a las empresas fabricantes de hardware a desarrollar supercomputadoras que consumen gigantescas cantidades de energía eléctrica y que producen tanto calor que requieren de grandes instalaciones refrigeradas que aseguren un correcto funcionamiento y agrega un alto costo de funcionamiento. En un mundo con recursos energéticos limitados y constante demanda por mayor poder computacional, el problema del consumo energético se presenta como uno de los mayores obstáculos para el diseño de sistemas que sean capaces de alcanzar la escala de los Exaflops. Por lo tanto, la comunidad científica está en la búsqueda de diferentes maneras de mejorar la eficiencia energética de los sistemas HPC [EVA10].

El cambio tecnológico, fundamentalmente a partir de los procesadores multicore, ha impuesto la necesidad de investigar en paradigmas "híbridos", en los cuales coexisten esquemas de memoria compartida con mensajes [LEI12]. Asimismo, la utilización de aceleradores combinados con las CPUs (GPU, FPGA, Xeon Phi) presenta una alternativa para alcanzar un alto speedup en determinadas clases de aplicaciones [KIN09][SIN12][JEF16]. Estos sistemas heterogéneos son capaces de obtener picos de rendimiento muy superiores a los de las CPUs, el incremento en el poder de cómputo se logra al mismo tiempo que se

limita el consumo de potencia energética [REI16]. Debe notarse que el modelo de programación orientado a estas arquitecturas cambia sensiblemente y la optimización de código paralelo requiere nuevos recursos.

En la actualidad también se comercializan placas de bajo costo como Raspberry PI [RAS16] u Odroid [ODR16] que poseen múltiples núcleos de baja complejidad y en algunos casos son procesadores multicore asimétricos (AMPs) que integran en un mismo chip diversos tipos de cores con distintas características (frecuencia, microarquitectura o consumo) pero con el mismo repertorio de instrucciones. Asimismo, existen diversos dispositivos móviles con capacidades similares. Es de interés estudiar como explotar el paralelismo en estos dispositivos para mejorar el rendimiento y/o consumo energético de las aplicaciones [ANN12], así como las características de scheduling en los mismos [POU17]. Por otra parte, los avances en las tecnologías de virtualización y cómputo distribuido han dado origen al paradigma de Cloud Computing [VEL09], que se presenta como alternativa a los tradicionales sistemas de Clusters y Multicluster para ambientes de HPC [ROD07], integrando grandes conjuntos de recursos virtuales (hardware, plataformas de desarrollo y/o servicios), fácilmente accesibles y utilizables como servicio ("as a service") por usuarios distribuidos por medio de una interfaz de administración web, con un modelo de arquitectura "virtualizada" [SHA10][XIN12]. Estos recursos pueden ser dinámicamente reconfigurados para adaptarse a una carga de trabajo variable (escalabilidad), permitiendo optimizar su uso evitando el sobre o sub dimensionamiento (elasticidad) [VAZ09].

Más allá de las potenciales características y beneficios que brinda un Cloud, de por sí atractivas, es de gran interés estudiar el despliegue de entornos de ejecución para cómputo paralelo y distribuido (Clusters Virtuales), como así también realizar I/D en la portabilidad de las aplicaciones de HPC en el Cloud [DOE11][ROD11]. Al enfocarse en Cloud Computing aparecen problemas clásicos de la Ciencia Informática, extendidos para este nuevo modelo de arquitectura: planificación, virtualización, asignación dinámica de recursos, migración de datos y procesos, y también su uso para aplicaciones de Big Data o las que requieren centralizar el accionar de "robots" distribuidos en tiempo real [MAY13]. Justamente Cloud Robotics es una de las áreas más prometedoras de la investigación informática actual en la cual se cuenta con "robots" dotados de diferentes sensores y capacidades, conectados a un Cloud vía Internet. Los temas de investigación derivados son múltiples: sensores, redes de sensores e inteligencia distribuida; robótica y sistemas colaborativos de tiempo real basados en robots; aplicaciones críticas (por ej. en ciudades inteligentes o en el ámbito industrial) [GUO12][KEH15][ROD16].

Los avances tecnológicos y la búsqueda continua de mayor eficiencia hacen necesario investigar diferentes componentes de las arquitecturas. En particular, las aplicaciones científicas con un uso intensivo de datos utilizan software de E/S paralelo para acceder a archivos [ENC15]. Contar con herramientas que permitan predecir el comportamiento de este tipo de aplicaciones en HPC es de gran utilidad para los desarrolladores. Para esto se puede utilizar el modelado basado en agentes y simulación (Agent-Based Modeling and Simulation, ABMS) [MAC06] [MEN13].

La importancia que ha adquirido el uso del paralelismo para aplicaciones científicas de gran duración requiere la necesidad de estudiar los problemas de detección y tolerancia a fallos en arquitecturas paralelas, debido al alto costo de relanzar la ejecución desde el comienzo en caso de resultados incorrectos. Esto se debe lograr tratando de minimizar el overhead temporal y de aprovechar la redundancia de recursos de hardware que caracteriza a estas arquitecturas [GOL09][FIA11]. El manejo de fallos es una preocupación creciente en HPC, se esperan crecimientos en las tasas de errores, mayores latencias de detección y elevadas cantidades de fallos silenciosos con capacidad de corromper los resultados de las aplicaciones. Por ello, se han desarrollado estrategias de detección y recuperación de fallos transitorios basadas en replicación de software, detectando divergencias en las comunicaciones entre réplicas para evitar que la corrupción se propague a otros procesos, restringiendo así la latencia de detección [MON17]. De esta forma se permiten obtener ejecuciones fiables con resultados correctos o conducir al sistema a una parada segura. La recuperación puede lograrse mediante múltiples checkpoints de nivel de sistema o de un único checkpoint de capa de aplicación.

Desde otro punto de vista, interesan también los problemas que significan integración de redes de sensores con modelos del mundo real (por ej. modelos meteorológicos, hídricos o de terreno) para prevención de emergencias [GAU16]. En esta línea, el eje del proyecto sigue estando en la problemática del paralelismo combinado con sistemas de tiempo real, pudiendo contribuir a proyectos multidisciplinarios, en particular por temas de emergencias hídricas, exploración de recursos naturales y temas de atención sanitaria y evacuación de edificios en situaciones de emergencia.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

- Estudio y caracterización de arquitecturas paralelas: clusters, grids, clouds, aceleradores (GPU, FPGA, Xeon Phi), placas de bajo costo (Raspberry PI, Odroid) e híbridos.
- Desarrollo de algoritmos de planificación de procesos orientado a procesadores asimétricos para

optimizar el rendimiento general. Análisis en los diferentes niveles: sistema operativo, compiladores, técnicas de programación.

- Desarrollo de aplicaciones concretas (numéricas y no numéricas) sobre diferentes máquinas paralelas utilizando técnicas de optimización adecuadas a cada arquitectura.
- Desarrollo de técnicas de tolerancia a fallas en sistemas paralelos y distribuidos, lo cual supone una mejora en el aprovechamiento de la redundancia de recursos que no resultan eficientemente utilizadas en dichas arquitecturas.
- Desarrollo de herramientas para la transformación de código heredado, buscando su optimización sobre arquitecturas paralelas.
- Integración de métricas de rendimiento computacional y energético. Predicción de performance de aplicaciones paralelas.
- Cloud Computing. Software de base. Desarrollo de aplicaciones de HPC (principalmente de big data).
- Sistemas inteligentes distribuidos de tiempo real aprovechando la potencia de cómputo del Cloud (Cloud Robotics).
- Utilización de ABMS para desarrollar un modelo de Entrada/Salida en HPC que permita predecir cómo cambios realizados en los diferentes componentes del mismo afectan a la funcionalidad y el rendimiento del sistema.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Estudiar modelos complejos, que integren redes de sensores en tiempo real y cómputo paralelo. Estrategias de predicción de catástrofes (inundaciones, incendios por ejemplo) se basan en estos modelos con alta capacidad de procesamiento y monitoreo de señales en tiempo real [GAU16].
- Se han desarrollado diferentes aplicaciones adaptadas para diferentes arquitecturas "híbridas" (que combinan clusters, multicores y aceleradores), y analizado/comparado el rendimiento obtenido [RUC16][POU16][MON16] [RUC17].
- Se está trabajando en técnicas de recuperación a partir de múltiples checkpoints de nivel de sistema, que sirvan para garantizar la correcta finalización de aplicaciones científicas sobre sistemas de HPC, que resultan afectadas por la ocurrencia de fallas transitorias externas y aleatorias, integrando esta solución con las herramientas de detección desarrolladas previamente [MON17].
- Se han desarrollado nuevos planificadores de tareas para multicores asimétricos sobre diferentes sistemas operativos con el objetivo de maximizar el rendimiento y minimizar el consumo de energía [SAE15][POU17].
- Desarrollo de un modelo de la Entrada/Salida en HPC por medio de ABMS (Agent-Based Modeling and Simulation) que permita predecir cómo cambios realizados en los diferentes componentes

del modelo afectan a la funcionalidad y el rendimiento del sistema [ENC15].

- Desarrollo de aplicaciones vinculadas con "Big Data", especialmente para resolver en Cloud Computing [BAS17].
- Optimización de algoritmos paralelos para controlar el comportamiento de múltiples robots que trabajan colaborativamente, considerando la distribución de su capacidad de procesamiento "local" y la coordinación con la potencia de cómputo y capacidad de almacenamiento (datos y conocimiento) de un Cloud.
- Actualización y modernización de código fuente de Sistemas Heredados (Legacy Systems) de Cómputo Científico a través de la aplicación de un proceso de desarrollo iterativo e incremental dirigido por transformaciones de código fuente, apoyado fuertemente en las herramientas de desarrollo. Dichas transformaciones se implementan para ser aplicadas automáticamente en un entorno integrado de desarrollo [MEN14][TIN15][MEN16].
- Trabajar en la implementación de transformaciones que ayuden a la paralelización del código fuente, así como también en herramientas de análisis estático de Código fuente [TIN13][TIN15].
- Adaptar las técnicas de scheduling y mapeo de procesos a procesadores de acuerdo a los objetivos actuales (en particular los relacionados con el consumo), considerando la migración dinámica de datos y procesos en función de rendimiento y consumo [GRA03][DEG10]. Se debe incluir la utilización de los registros de hardware de los procesadores para la toma de diferentes decisiones en tiempo de ejecución.
- Analizar metodologías y herramientas de software y hardware para medir consumo energético. Determinar el grado de error en las herramientas de medición por software. Determinar el consumo energético de cada componente de las máquinas (memoria, disco, procesador, etc) [PAN18][PIP17].

4. FORMACIÓN DE RECURSOS HUMANOS

En cooperación con Universidades iberoamericanas se ha implementado la Maestría en Cómputo de Altas Prestaciones y se continúa dictando la Especialización en Cómputo de altas Prestaciones y Tecnología GRID. Asimismo, se tiene un importante número de doctorandos (del país y del exterior) realizando el Doctorado en Ciencias Informáticas de la UNLP.

Desde el año 2013 se organizan anualmente las Jornadas de Cloud Computing & Big Data (JCC&BD), integrando una Escuela con cursos de Posgrado relacionados con la temática de las líneas de investigación presentadas.

Existe cooperación a nivel nacional e internacional y dentro de la temática del proyecto se espera alcanzar 5

Tesis de Doctorado y 5 Tesis de Maestría en los próximos 3 años, en el país. Al menos tener 3 Doctorandos en el exterior/mixtos en el mismo período.

En 2017 se aprobó 1 Tesis Doctoral [POU17]. También se aprobaron 2 trabajos de Especialista y 2 Tesinas de grado.

5. BIBLIOGRAFÍA

[ANN12] Annamalai A., Rodrigues R., Koren I., Kundu S., "Dynamic Thread Scheduling in Asymmetric Multicores to Maximize Performance-per-Watt," 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, pp. 964-971, 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, 2012.

[BAS17] Basgall, M. J., Hasperué, W., Naiouf, M., & Bariviera, A. F. "Cálculo del exponente de Hurst utilizando Spark Streaming: enfoque experimental sobre un flujo de transacciones de criptomonedas. Presentado en XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017).

[DEG10] De Giusti L., Chichizola F., Naiouf M., De Giusti A.E., Luque E. "Automatic Mapping Tasks to Cores - Evaluating AMTHA Algorithm in Multicore Architectures". IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 2, No 1, March 2010. ISSN (Online): 1694-0784. ISSN (Print): 1694-0814. Págs. 1-6.

[DOE11] Doelitzcher, F., Held, M., Sulistio, A., Reich, C. ViteraaS: "Virtual Cluster as a Service". In: 3rd IEEE International Conference on Cloud Computing Technology and Science. Grecia (2011).

[ENC15] D. Encinas et al., "Modeling I/O System in HPC: An ABMS Approach". The Seventh International Conference on Advances in System Simulation (SIMUL), ISBN: 978-1-61208-442-8, 2015.

[EVA10] J. Evans, "On performance and energy management in high performance computing systems" in Parallel Processing Workshops (ICPPW). 39th International Conference on, pp. 445-452. 2010.

[FIA11] Fialho L. "Fault Tolerance configuration for uncoordinated checkpoints". Ph.D. Thesis, Universidad Autónoma de Barcelona, Barcelona, España, Julio 2011.

[GAU16] Adriana Gaudiani, Emilio Luque, Pablo García, Mariano Re, Marcelo Naiouf, Armando De Giusti. "How a Computational Method Can Help to Improve the Quality of River Flood Prediction by Simulation". Advances and New Trends in Environmental and Energy Informatics (part V). ISBN 978-3-319-23455-7. Pp337-351. 2016.

[GOL09] Golander A., Weiss S., Ronen R. "Synchronizing Redundant Cores in a Dynamic DMR Multicore Architecture". IEEE Transactions on Circuits and Systems II: Express Briefs Volume 56, Issue 6, 474-478. 2009.

- [GRA03] Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing". Second Edition. Pearson Addison Wesley, 2003.
- [GUO12] Guoqiang Hu, Wee Peng Tay, Yonggang Wen: "Cloud robotics: architecture, challenges and applications". In: Network, IEEE, vol.26, no.3, pp.21-28. 2012.
- [JEF16] Jeffers, James; Reinders, James; Sodani, Avinash. "Intel Xeon Phi Processor High Performance Programming. Knights Landing Edition", Morgan Kaufmann, 2016.
- [KEH15] Kehoe B., Patil S., Abbeel P., Goldberg K.: "A Survey of Research on Cloud Robotics and Automation". In: IEEE Transactions on Automation Science and Engineering (T-ASE): Special Issue on Cloud Robotics and Automation. Vol. 12, no. 2. 2015.
- [KIN09] Kindratenko, V.V.; Enos, J.J.; Guochun Shi; Showerman, M.T.; Arnold, G.W.; Stone, J.E.; Phillips, J.C.; Wen-Mei Hwu, "GPU clusters for high-performance computing," Cluster Computing and Workshops, 2009. CLUSTER '09. IEEE International Conference on, vol., no., pp.1,8, Aug. 31 2009-Sept. 4 2009
- [KIR12] Kirk D., Hwu W. "Programming Massively Parallel Processors, second edition: A Hands-on Approach. Morgan-Kaufmann. 2012.
- [LEI12] Leibovich F., Chichizola F., De Giusti L., Naiouf M., Tirado Fernández F., De Giusti A. "Programación híbrida en clusters de multicore. Análisis del impacto de la jerarquía de memoria". XII Workshop de Procesamiento Distribuido y Paralelo. CACIC2012. ISBN: 978987-1648-34-4. Pág. 306-315. 2012.
- [MAC06] C. Macal, M. North, "Tutorial on agent-based modeling and simulation part 2: how to model with agents", in: Proceedings of the Winter Simulation Conference, 2006.
- [MAY13] Mayer-Schönberger V., Cukier K., "Big Data: A revolution that will transform how we live, work and think" Houghton Mifflin Harcourt 2013.
- [MCC12] McCool M., Robison A., Reinders J. "Structured Parallel Programming" Elsevier-Morgan Kaufmann, 2012.
- [MEN13] Mendez S., "Metodología para la evaluación de prestaciones del Sistema de Entrada/Salida en computadoras de altas prestaciones". Tesis de Doctorado en Computación de Altas Prestaciones. Universidad Autónoma de Barcelona, 2013.
- [MEN14] Méndez M., Tinetti F. "Integrating Software Metrics for Fortran Legacy into an IDE". XI Workshop Ingeniería de Software – CACIC 2014. ISBN 978-987-3806-05-6. Pág. 771-780. San Justo, Buenos Aires, Argentina. Octubre 2014.
- [MEN16] Méndez M. "Aplicaciones de Cómputo Científico: Mantenimiento del Software Heredado". Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016
- [MON16] E. Montes de Oca, L. De Giusti, F. Chichizola, A. De Giusti, and M. Naiouf. "Análisis de uso de un algoritmo de balanceo de carga estático en un Cluster Multi-GPU Heterogéneo". Actas del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016), ISBN: 978-987-733-072-4, págs. 159-168. 2016.
- [MON17] D. M. Montezanti, A. E. De Giusti, M. Naiouf, J. Villamayor, D. Rexachs del Rosario, and E. Luque. "A Methodology for Soft Errors Detection and Automatic Recovery". Proceedings of the 2017 International Conference on High Performance Computing Simulation (HPCS2017), págs. 434-441. 2017.
- [ODR16] Odroid <http://www.hardkernel.com> Accedido 21 de Marzo de 2016.
- [PAN18] J. M. Paniego, S. Gallo, M. P. Puig, F. Chichizola, L. D. Giusti, and J. Ballardini. "Analysis of RAPL Energy Prediction Accuracy in a Matrix Multiplication Application on Shared Memory". Computer Science – CACIC 2017. Communications in Computer and Information Science, vol 790, ISBN: 978-3-319-75213-6 978-3-319-75214-3, Springer, Cham, págs. 37-46, 2018.
- [PIP17] M. Pi Puig, L. C. De Giusti, M. Naiouf, A. E. De Giusti. "GPU Performance and Power Consumption Analysis: A DCT based denoising application". Actas del XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017). ISBN: 978-950-34-1539-9, págs. 185-195. 2017.
- [POU16] Pousa A, Sanz V, De Giusti A. "Estructurando código paralelo para clusters heterogéneos de CPUs/GPUs". Proceedings del XXII Congreso Argentino de Ciencias de la Computación. ISBN: 978-987-733-072-4. Pp. 139-148
- [POU17] A. Pousa, Directores: De Giusti, Armando; Saez Alcaide, Juan Carlos. "Optimización de rendimiento, justicia y consumo energético en sistemas multicore asimétricos mediante planificación". Tesis de Doctorado en Ciencias Informáticas. Facultad de Informática (UNLP). 2017.
- [RAS16] Raspberry PI. <https://www.raspberrypi.org/> Accedido 21 de Marzo de 2016.
- [RAU10] Rauber T., Rüniger G. "Parallel programming for multicore and cluster systems". Springer. 2010.
- [REI16] Reinders, J., Jeffers, J., Sodani, A. "Intel Xeon Phi Processor High Performance Programming Knights Landing Edition". Morgan Kaufmann Publishers Inc., Boston, MA, USA, 2016
- [ROD07] Rodríguez, I. P., Pousa, A., Pettoruti, J.E., Chichizola, F., Naiouf, M., De Giusti, L., De Giusti, A.: "Estudio del overhead en la migración de algoritmos paralelos de cluster y multicluster a GRID". In: Proceedings del XIII Congreso Argentino de Ciencias de la Computación. Argentina (2007).
- [ROD11] Rodríguez, I., Pettoruti, J.E., Chichizola, F., De Giusti, A.: "Despliegue de un Cloud Privado para entornos de cómputo científico". In: Proceedings del XVII Congreso Argentino de Ciencias de la Computación. Argentina (2011).

- [ROD16] Rodriguez, I., Paniego, J. M., Rodriguez Eguren, S., Estrebou, C., De Giusti, A. "Cloud Robotics: Sistema Multi-Robot conectado al Cloud público AWS". Actas del XXII Congreso Argentino de Ciencias de la Computación. Argentina (CACIC 2016). Págs. 189-198.
- [RUC16] Rucci, Enzo. "Evaluación de rendimiento y eficiencia energética en sistemas heterogéneos para bioinformática". Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016.
- [RUC17] E. Rucci, A. E. De Giusti, and M. Naiouf. "Blocked All-Pairs Shortest Paths Algorithm on Intel Xeon Phi KNL Processor: A Case Study". Actas del XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017)., ISBN: 978-950-34-1539-9, págs. 154-164. 2017.
- [SAE15] Juan Carlos Saez, Adrian Pousa, Daniel Chaver, Fernando Castro, Manuel Prieto Matias. ACFS: A Completely Fair Scheduler for Asymmetric Single-ISA Multicore Systems. ACM SAC 2015 (The 30TH ACM/SIGAPP Symposium on applied computing). 2015.
- [SHA10] Shafer, J.: "I/O virtualization bottlenecks in cloud computing today". In: Proceedings of the 2nd conference on I/O virtualization (VIOV10). USA (2010).
- [SIN12] Sinha, R.; Prakash, A.; Patel, H.D., "Parallel simulation of mixed-abstraction SystemC models on GPUs and multicore CPUs," Design Automation Conference (ASP-DAC), 2012 17th Asia and South Pacific, pp.455,460. 2012.
- [TIN13] F. G. Tinetti, M. Méndez, A. De Giusti. "Restructuring Fortran legacy applications for parallel computing in multiprocessors". The Journal of Supercomputing, May 2013, Volume 64, Issue 2, ISSN 0920-8542, DOI 10.1007/s11227-012-0863-x, pp 638-659.
- [TIN15] M. Mendez, F. G. Tinetti. "Integrating Software Metrics for Fortran Legacy into an IDE" Computer Science & Technology Series – XX Argentine Congress of Computer Science, Selected Papers., ISBN: 978-987-1985-71-5, Red UNCI, págs. 126-134, 2015.
- [VAZ09] Vázquez Blanco C., Huedo E., Montero R. S., Llorente I. M. "Elastic Management of Cluster-based Services in the Cloud". Proceedings pp. 19-24, ACM Digital Library 2009. ISBN 978-1-60558-564-2.
- [VEL09] Velte, A.T., Velte, T.J., Elsenpeter, R.: "Cloud Computing: A Practical Approach". McGraw Hill Professional. 2009.
- [XIN12] Xing, Y., Zhan, Y.: "Virtualization and Cloud Computing". In: Proceedings pp.305-312, Springer Link. ISBN 978-3-642-27323-0. (2012). Morgan Kaufmann. 2013.

CÓMPUTO PARALELO Y DISTRIBUIDO PARA HPC. FUNDAMENTOS, CONSTRUCCIÓN Y EVALUACIÓN DE APLICACIONES.

Marcelo Naiouf⁽¹⁾, Armando De Giusti⁽¹⁾⁽²⁾, Laura De Giusti⁽¹⁾⁽³⁾, Franco Chichizola⁽¹⁾, Victoria Sanz⁽¹⁾⁽²⁾⁽³⁾, Adrián Pousa⁽¹⁾, Enzo Rucci⁽¹⁾⁽²⁾, Silvana Gallo⁽¹⁾⁽²⁾, Erica Montes de Oca⁽¹⁾, Emmanuel Frati⁽¹⁾, Mariano Sánchez⁽¹⁾, María José Basgall⁽¹⁾⁽²⁾, Adriana Gaudiani⁽⁴⁾

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 115, La Plata, Buenos Aires
Comisión de Investigaciones Científicas de la Pcia. de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CIC – Comisión de Investigación Científica de la Provincia de Buenos Aires

⁴Universidad Nacional de General Sarmiento

{mnaiouf, degiusti, ldgiusti, francoch, vsanz, apousa, erucci, sgallo, emontesdeoca, fefrati, msanchez, mjbassgall}@lidi.info.unlp.edu.ar, agaudi@ungs.edu.ar

RESUMEN

El eje central de la línea son los temas de procesamiento paralelo y distribuido para HPC (fundamentos y aplicaciones). Interesa la construcción, evaluación y optimización de soluciones sobre diferentes plataformas de software y arquitecturas con múltiples procesadores (multicore, clusters, cloud, aceleradores y placas de bajo costo), los lenguajes y paradigmas de programación paralela (puros e híbridos), los modelos de representación de aplicaciones paralelas, los algoritmos de mapping y scheduling, el balance de carga, las métricas de evaluación de complejidad y rendimiento computacional y energético, y la construcción de ambientes para la enseñanza de la programación concurrente y paralela.

Se propone aplicar los conceptos en problemas numéricos y no numéricos de cómputo intensivo y/o sobre grandes volúmenes de datos (búsquedas, simulaciones, n-body, big data, reconocimiento de patrones, bioinformática, etc), con el fin de obtener soluciones de alto rendimiento.

En la dirección de tesis de postgrado existe colaboración con el grupo HPC4EAS (High Performance Computing for Efficient Applications and Simulation) del Dpto. de Arquitectura de Computadores y Sistemas Operativos de la Universidad Autónoma de Barcelona, y con el Departamento de Arquitectura de Computadores y Automática de la Universidad Complutense de Madrid, entre otros.

Palabras clave: Cómputo paralelo y distribuido de altas prestaciones. Algoritmos paralelos y distribuidos. Clusters. Multicore. Aceleradores. Consumo energético. Balance de carga. Aplicaciones. Performance.

CONTEXTO

La línea de I/D que se presenta es parte del Proyecto “Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real” del III-LIDI acreditado por el Ministerio de Educación, del proyecto “Computación de Alto Desempeño, Minería de Datos y Aplicaciones de interés social en la Provincia de

Buenos Aires” financiado por la CIC PBA en la convocatoria a Proyectos de Innovación y Transferencia en Areas Prioritarias de la Pcia. de Buenos Aires (PIT-AP-BA), y de proyectos acreditados y subsidiados por la Facultad de Informática de la UNLP. Además, existe cooperación con Universidades de Argentina, Latinoamérica y Europa a través de proyectos acreditados por AECID, CyTeD, OEI y CIC y becas de Telefónica de Argentina. Asimismo, el III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD).

1. INTRODUCCIÓN

El área de cómputo de altas prestaciones (HPC, High-Performance Computing) es clave dentro de las Ciencias de la Computación, debido al creciente interés por el desarrollo de soluciones a problemas con alta demanda computacional y de almacenamiento, produciendo transformaciones profundas en las líneas de I/D [GIL14]. El rendimiento en este caso está relacionado con dos aspectos: las arquitecturas de soporte y los algoritmos que hacen uso de las mismas, y el desafío se centra en cómo aprovechar las prestaciones obtenidas a partir de la evolución de las arquitecturas físicas. En esta línea la mayor importancia está en los algoritmos paralelos y en los métodos utilizados para su construcción y análisis a fin de optimizarlos.

Uno de los cambios de mayor impacto ha sido el uso de manera masiva de procesadores con más de un núcleo (*multicore*), produciendo plataformas distribuidas híbridas (memoria compartida y distribuida) y generando la necesidad de desarrollar sistemas operativos, lenguajes y algoritmos que las usen adecuadamente. También creció la incorporación de placas aceleradoras a los sistemas multicore constituyendo plataformas paralelas de memoria compartida con paradigma de programación propio asociado como pueden ser las unidades de procesamiento gráfico (GPU, Graphic Processing Unit) de NVIDIA y AMD, los coprocesadores Xeon Phi de Intel [JEF13] o los aceleradores basados en circuitos integrados reconfigurables (FPGAs, Field Programmable Gate Array) [SET13]. En la actualidad se comercializan placas

de bajo costo como Raspberry PI [RAS16] u Odroid [ODR16] que poseen múltiples núcleos de baja complejidad y en algunos casos son procesadores multicore asimétricos (AMPs) con el mismo repertorio de instrucciones. Es de interés estudiar como explotar el paralelismo en estos dispositivos para mejorar el rendimiento y/o consumo energético de las aplicaciones [ANN12], así como las características de scheduling en los mismos [SAE15]. Asimismo, los entornos de computación cloud introducen un nuevo foco desde el punto de vista del HPC, brindando un soporte “a medida” sin la necesidad de adquirir el hardware.

La creación de algoritmos paralelos en arquitecturas multiprocesador no es un proceso directo [MCC12]. El costo puede ser alto en términos del esfuerzo de programación y el manejo de la concurrencia adquiere un rol central en el desarrollo. Si bien en las primeras etapas el diseñador de una aplicación paralla puede abstraerse de la máquina sobre la que ejecutará el algoritmo, para obtener buen rendimiento debe tenerse en cuenta la plataforma de destino. En las máquinas multiprocesador, se deben identificar las capacidades de procesamiento, interconexión, sincronización y escalabilidad. La caracterización y estudio de rendimiento del sistema de comunicaciones es de interés para la predicción y optimización de performance, así como la homogeneidad o heterogeneidad de los procesadores [DEG10].

Muchos problemas algorítmicos se vieron impactados por los multicore y clusters de multicore. A partir de incorporar varios chips multicore dentro de un nodo y conectar múltiples nodos vía red, se puede crear una arquitectura NUMA, de modo que los cores en un chip compartan memoria principal, y puedan acceder remotamente a la memoria dedicada de otro chip, aunque ese acceso sea más costoso, surgiendo así varios niveles de comunicación. Esto impacta sobre el desarrollo de algoritmos que aprovechen adecuadamente las arquitecturas, y motiva el estudio de performance en sistemas híbridos. Además, es necesario estudiar la utilización de diferentes lenguajes y bibliotecas ya que aún no se cuenta con un standard, aunque puede mencionarse el uso de los tradicionales MPI, OpenMP y Pthreads [CHA08][HAG11] o los más recientemente explorados UPC, Chapel y Titanium del modelo PGAS [DEW15].

La combinación de arquitecturas con múltiples núcleos con aceleradores dio lugar a plataformas híbridas con diferentes características. Más allá del tipo de acelerador utilizado, la programación de esta clase de plataformas representa un verdadero desafío. Para lograr aplicaciones de alto rendimiento, los programadores deben enfrentar dificultades como: estudiar características específicas de cada arquitectura y aplicar técnicas de programación y optimización particulares para cada una de ellas, lograr un balance de carga adecuado entre los diferentes dispositivos de procesamiento y afrontar la ausencia de estándares y para este tipo de sistemas.

Por otra parte, los avances en las tecnologías de virtualización han llevado a que Cloud Computing sea una alternativa a los tradicionales sistemas de cluster [EC213]. El uso de cloud para HPC presenta desafíos atractivos, brindando un entorno reconfigurable dinámicamente sin la necesidad de adquirir hardware, y es una excelente plataforma para testear escalabilidad de algoritmos aunque queda mucho por hacer en cuanto al diseño, lenguajes y programación

Métricas de evaluación del rendimiento y balance de carga

La diversidad de opciones vuelve complejo el análisis de performance de los Sistemas Paralelos, ya que los ejes sobre los cuales pueden compararse dos sistemas son varios. Existe un gran número de métricas para evaluar el rendimiento, siendo las tradicionales: tiempo de ejecución, speedup, eficiencia.

Por su parte, la *escalabilidad* permite capturar características de un algoritmo paralelo y la arquitectura en que se lo implementa. Posibilita testear la performance de un programa sobre pocos procesadores y predecirla en un número mayor, así como caracterizar la cantidad de paralelismo inherente en un algoritmo.

Un aspecto de interés que se ha sumado como métrica, a partir de las plataformas con gran cantidad de procesadores, es el del consumo y la eficiencia energética [BAL13]. Muchos esfuerzos están orientados a tratar el consumo como eje de I/D, como métrica de evaluación, y también a la necesidad de metodologías para medirlo.

El objetivo principal del cómputo paralelo es reducir el tiempo de ejecución haciendo uso eficiente de los recursos. El *balance de carga* es un aspecto central y consiste en, dado un conjunto de tareas que comprenden un algoritmo y un conjunto de procesadores, encontrar el mapeo (asignación) de tareas a procesadores tal que cada una tenga una cantidad de trabajo que demande aproximadamente el mismo tiempo, y esto es más complejo si hay heterogeneidad. Dado que el problema general de mapping es *NP-completo*, pueden usarse enfoques que dan soluciones subóptimas aceptables. Las técnicas de planificación a nivel micro (dentro de cada procesador) y macro (en un cluster) deben ser capaces de obtener buen balance de carga. Existen técnicas estáticas y dinámicas cuyo uso depende del conocimiento que se tenga sobre las tareas que componen la aplicación. Entre las aplicaciones de interés se encuentran las numéricas y no numéricas con alta demanda de cómputo.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

- Investigar en temas de cómputo paralelo y distribuido de alto desempeño, tanto en lo referido a los fundamentos como a la construcción y evaluación de las aplicaciones. Esto incluye los problemas de software asociados con el uso de arquitecturas multiprocesador:

- Lenguajes, modelos y paradigmas de programación paralela (puros e híbridos a distintos niveles).
- Asignación de procesos a procesadores optimizando el balance de la carga de procesamiento.
- Métricas de evaluación de complejidad y rendimiento: speedup, eficiencia, escalabilidad, consumo energético, costo de programación.
- Construir, evaluar y optimizar soluciones utilizando algoritmos concurrentes, paralelos y distribuidos sobre diferentes plataformas de software y arquitecturas con múltiples procesadores:
 - Arquitecturas de trabajo homogéneas, heterogéneas e híbridas: multicores, clusters, GPU, Xeon Phi, FPGA, placas de bajo costo y cloud.
 - Se propone aplicar los conceptos en problemas numéricos y no numéricos de cómputo intensivo y/o sobre grandes volúmenes de datos (aplicaciones científicas, búsquedas, simulaciones, imágenes, realidad virtual y aumentada, bioinformática, big data, n-body).
- Analizar y desarrollar ambientes para la enseñanza de programación concurrente y paralela.
 - Caracterizar modelos de arquitecturas paralelas.
 - Representar distintos modelos de comunicación/sincronización.
 - Definir métricas de evaluación de rendimiento y eficiencia energética.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Desarrollar y optimizar algoritmos paralelos sobre diferentes modelos de arquitectura. En particular, en aplicaciones numéricas y no numéricas de cómputo intensivo y tratamiento de grandes volúmenes de datos (big data).
- Utilizar arquitecturas híbridas que combinan memoria compartida y pasaje de mensajes, evaluando performance para distintos modelos de comunicación.
- Estudiar y comparar los lenguajes sobre las plataformas multiprocesador para diferentes modelos de interacción entre procesos.
- Investigar la paralelización en plataformas que combinan clusters, multicore y aceleradores. Comparar estrategias de distribución de trabajo teniendo en cuenta las diferencias en potencias de cómputo y comunicación, dependencia de datos y memoria requerida.
- Evaluar la performance (speedup, eficiencia, escalabilidad, consumo energético) de las soluciones propuestas. Analizar el rendimiento de soluciones paralelas a problemas con diferentes características (dependencia de datos, relación cómputo / comunicación, memoria requerida).
- Mejorar y adecuar las técnicas disponibles para el balance de carga (estático y dinámico) entre procesos a las arquitecturas consideradas.

En este marco, pueden mencionarse los siguientes resultados:

- Para la experimentación se han utilizado y analizado diferentes arquitecturas homogéneas o heterogéneas, incluyendo multicores, cluster de multicores (con 128 núcleos), GPU y cluster de GPU, Xeon Phi y FPGA.
- Se experimentó la paralelización en arquitecturas híbridas, con el objetivo de estudiar el impacto del mapeo de datos y procesos, así como de los lenguajes y librerías utilizadas.
- Respecto de las aplicaciones estudiadas y algoritmos implementados, se trabajó fundamentalmente con los siguientes problemas:
 - **Best-first search (BFS) paralelo.** El algoritmo BFS es utilizado para resolver problemas combinatorios, los cuales requieren encontrar una secuencia de acciones que transformen una configuración inicial (problema) en una configuración final (solución). En particular, A* es una variante de BFS que permite encontrar soluciones de costo óptimo. Estos algoritmos requieren una alta capacidad de cómputo y gran cantidad de memoria, por esto su paralelización es imprescindible. En los últimos años se ha re-impulsado el desarrollo de algoritmos paralelos BFS para aprovechar: (a) la potencia de cómputo de los procesadores *multicore* (b) la gran cantidad de RAM y potencia de cómputo de los *clusters de multicore*. En este sentido, HDA* [KIS13] paraleliza A* sobre clusters utilizando MPI: cada procesador realiza una búsqueda cuasi-independiente y se distribuyen los nodos en base a una función hash estándar. Otros autores [BUR10] adaptaron HDA* a máquinas multicore utilizando Pthreads: eliminan overheads existentes en la versión original al correr sobre una arquitectura de memoria compartida y utilizan menor cantidad de memoria. Para explotar eficientemente los recursos de un *cluster de multicore*, desarrollamos Hybrid HDA* (HHDA*) [SAN17], una versión híbrida de HDA* programada con MPI+Pthreads. El trabajo experimental demostró que HHDA* alcanza un rendimiento superior y consume menor cantidad de memoria, comparado con HDA* (versión original). Estas mejoras permitieron a HHDA* resolver una de las instancias más complejas del caso de estudio. Como trabajo futuro planeamos paralelizar algoritmos de búsqueda *sub-óptimos* usando nuestra estrategia de paralelización híbrida.

- **Aceleración de aplicaciones con cómputo colaborativo CPU-GPU.** Las computadoras comerciales actuales incluyen decenas de cores y al menos una GPU. El uso de ambas unidades de procesamiento de forma colaborativa puede mejorar significativamente el rendimiento de una aplicación. Sin embargo, esto supone un desafío para los programadores ya que dichas unidades difieren en arquitectura, modelo de programación y rendimiento. Para facilitar la etapa de desarrollo, en [POU16]

propusimos un esquema para estructurar código paralelo a ser ejecutado sobre un cluster heterogéneo de CPUs/GPUs, utilizando todos los recursos disponibles (CPUs y GPUs). Aplicamos este esquema al problema de suma por reducción y comprobamos que es posible incrementar el rendimiento de la aplicación utilizando CPUs y GPUs en forma colaborativa. Planeamos aplicar nuestro esquema de cómputo colaborativo a distintos tipos de problemas demandantes en cómputo.

➤ **Alineamiento de secuencias biológicas.** Esta operación consiste en comparar dos o más secuencias biológicas, como pueden ser las de ADN o las de proteínas, y resulta fundamental en investigaciones de la bioinformática y la biología molecular. El algoritmo de Smith-Waterman es considerado el método de alineamiento más preciso. Desafortunadamente, este algoritmo resulta costoso debido a su complejidad computacional cuadrática mientras que la situación se agrava aún más a causa del crecimiento exponencial de datos biológicos en los últimos años [RUC16]. El reciente surgimiento de aceleradores en HPC (GPU, Xeon Phi, FPGA, entre otros) da la oportunidad de acelerar los alineamientos sobre hardware comúnmente disponible a un costo accesible. En primer lugar, se exploró el empleo del modelo de programación OpenCL sobre FPGAs para acelerar el alineamiento de secuencias largas de ADN [RUC17a]. Luego, se desarrolló una herramienta capaz de procesar alineamientos de secuencias de ADN sin restricciones de tamaño y se analizó su rendimiento comparándolo con el de otras implementaciones basadas en multicores, Xeon Phi y GPUs [RUC18a]. Por otra parte, la salida al mercado de la segunda generación de procesadores Xeon Phi (Knights Landing) dio la oportunidad de evaluar su uso para búsquedas de similitud en bases de datos de proteínas [RUC17b]. A futuro, interesa explorar el uso de próximas generaciones de procesadores y aceleradores para esta aplicación, como pueden ser los procesadores Xeon *Skylake* de Intel y las nuevas generaciones de FPGAs.

➤ **Cálculo de los caminos mínimos.** Es uno de los problemas básicos y de mayor antigüedad de la teoría de grafos teniendo aplicación en el dominio de las comunicaciones, del ruteo de tráfico, de la bioinformática, entre otros. El algoritmo de Floyd-Warshall (FW) permite computar la distancia mínima entre todos los pares de un grafo. Además de poseer una alta demanda de ancho de banda, FW resulta costoso computacionalmente al ser $O(n^3)$. Empezando por una implementación secuencial de base, se estudió y cuantificó cómo diferentes optimizaciones a nivel de datos, hilos y compilador permiten mejorar su rendimiento sobre los nuevos procesadores Xeon Phi Knights Landing [RUC18b]. Como trabajo futuro, interesa realizar un estudio comparativo con otros aceleradores (como GPUs), no sólo desde el punto de

vista del rendimiento sino también de la eficiencia energética.

➤ **Simulación distribuida de modelos orientados al individuo.** Debido al incremento en la complejidad de los modelos es necesaria mayor cantidad de cómputo y comunicación para lograr resultados representativos. Actualmente, se busca analizar el consumo energético del simulador teniendo en cuenta el balance de cómputo, ya que al trabajar con grandes cantidades de individuos se producen desbalances en cuanto a la cantidad de trabajo de cada uno de los diferentes procesos lógicos. Se pretende desarrollar un modelo energético para estimar el consumo de energía para este tipo de algoritmos y que permita decidir el escenario adecuado.

➤ **Problemas de tipo N-body.** Se utilizaron las plataformas de memoria compartida GPU y cluster de multicore para la resolución de problemas con alta demanda computacional del tipo N-body. Se emplearon diferentes modelos de comunicación: memoria compartida (Pthreads en CPU y CUDA en GPU), pasaje de mensajes (MPI) y soluciones híbridas (MPI-Pthreads). Se han mostrado los beneficios del uso de la GPU en problemas con características similares al caso planteado. El trabajo experimental ha dado como resultado una buena aceleración obtenida utilizando cluster de GPU. Además, se observó claramente que el uso del cluster de GPU logró una aceleración proporcional al speedup conseguido con el cluster de CPU pero con tiempos de ejecución significativamente menores [MON14]. También se han desarrollado diferentes alternativas de distribución de trabajo usando un Cluster de GPUs heterogéneas [MON16]. Además del uso de un Cluster de GPU se está utilizando MultiGPU, pero haciendo énfasis en el consumo energético. El énfasis en las experimentaciones actuales, se centra en la determinación de un modelo de estimación de consumo energético, y la precisión de la medición de los contadores de hardware en GPU.

➤ **Problemas de simulación relacionados con fenómenos naturales (inundaciones).** Análisis de diferentes soluciones para la paralelización de este tipo de aplicaciones que son intensivas en cómputo; y el tiempo de ejecución y la performance alcanzable son críticas dado que los resultados que se esperan determinarán alertas y toma de decisiones. La utilización de escenarios de simulación en entornos donde interesa estudiar el comportamiento en situaciones de desastres producidos por fenómenos naturales como las inundaciones. En este ámbito se avanza en dos temas: (1) La implementación de un método de sintonización de un simulador de inundaciones en ríos de llanura, mediante la técnica de simulación paramétrica. El proceso requiere lanzar miles de escenarios de simulación hasta encontrar un conjunto ajustado de parámetros de entrada del

simulador. La experimentación se lleva a cabo con un modelo master-worker sobre un cluster [GAU15]. (2) En colaboración con el Laboratorio de Hidrología de la UNLP se comenzó con la paralelización de la simulación de inundaciones producidas por lluvias (en particular en el ámbito de la ciudad de La Plata, donde una corrida "standard" es del orden de las 8 hs), a fin de reducir el tiempo de ejecución a pocos minutos y permitir establecer un sistema de alertas [GAU16].

➤ **Ambientes para la enseñanza de concurrencia.** Se desarrolló el entorno CMRE para la enseñanza de programación concurrente y paralela a partir de cursos iniciales en carreras de Informática. Incluye un entorno visual que representa una ciudad en la que pueden definirse varios robots que interactúan. Combina aspectos de memoria compartida y distribuida mediante instrucciones para bloquear y liberar esquinas de la ciudad y el concepto de pasaje de mensajes a través de primitivas de envío y recepción. Además, se incluyen los conceptos de heterogeneidad (diferentes velocidades de los robots) y consumo energético [DEG17]. Se ha integrado con el uso de robots físicos (Lego Mindstorm 3.0) que ejecutan en tiempo real las mismas instrucciones que los robots virtuales y se comunican con el entorno mediante bluetooth [DEG16]. Se ha ampliado para incorporar conceptos básicos de computación en la nube (Cloud Computing) [DEG16].

➤ **Aplicaciones en Big Data.** Para trabajar con Big Data, escenario que es cada vez más común debido al gran volumen de datos que se genera a diario, se requiere de cómputos de altas prestaciones y de herramientas específicas que ayuden al proceso de adaptación de los algoritmos iterativos tradicionales para ser ejecutados de manera paralela y distribuida. Una de las herramientas más utilizadas es Apache Spark, que se caracteriza principalmente por su velocidad de procesamiento debido a que hace uso intensivo de memoria RAM evitando continuas escrituras en disco y, a su vez, por poseer mecanismos eficientes de tolerancia a fallos. Apache Spark posee un módulo para el tratamiento de datos provenientes de un flujo de información potencialmente infinito llamado, Spark Streaming. En esta línea se está trabajando, por un lado, en la implementación de cálculos de índices económicos sobre un flujo de datos proveniente del mercado de las criptomonedas (Bitcoin, Ethereum, Litecoin, etc) utilizando Spark Streaming [BAR17] [BAS17]. Por otro lado, en la adaptación de técnicas tradicionales de Machine Learning para ser aplicadas a grandes volúmenes de datos mediante el uso del framework Spark.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de la temática de la línea de I/D en el año 2017 se concluyó 1 tesis doctoral, 1 Trabajo Final de Especialización y 1 Tesina de Grado de Licenciatura. Se

encuentran en curso en el marco del proyecto 3 tesis doctorales, 1 de maestría, 3 trabajos de Especialización y 2 Tesinas.

Se participa en el dictado de las carreras de Doctorado en Cs. Informáticas y Magister y Especialización en Cómputo de Altas Prestaciones de la Facultad de Informática UNLP, por lo que potencialmente pueden generarse más Tesis y Trabajos Finales.

Hay cooperación con grupos de otras Universidades del país y del exterior, y tesis de diferentes Universidades realizan su trabajo con el equipo del proyecto.

5. BIBLIOGRAFÍA

[ANN12] Annamalai A., Rodrigues R., Koren I., Kundu S., "Dynamic Thread Scheduling in Asymmetric Multicores to Maximize Performance-per-Watt," 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, pp. 964-971, 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, 2012.

[BAL13] Ballardini J., Rucci E., De Giusti A., Naiouf M., Suppi R., Rexachs D., Luque E. "Power Characterisation of Shared-Memory HPC Systems". Computer Science & Technology Series – XVIII Argentine Congress of Computer Science Selected Papers. ISBN 978-987-1985-20-3. Pp. 53-65. EDULP, La Plata (Argentina), 2013

[BAR17] Bariviera, A. F., Basgall, M. J., Hasperué, W., & Naiouf, M. "Some stylized facts of the Bitcoin market". Physica A: Statistical Mechanics and its Applications, 484, 82790. 2017.

[BUR10] Burns E, Lemons S, Ruml W, Zhou R. "Best First Heuristic Search for Multicore Machines". Journal of Artificial Intelligence Research, Vol.39, No.1, pp. 689-743, 2010.

[CHA08] Chapman, B., Jost, G. & Van der Pas. Using OpenMP – Portable Shared Memory Parallel Programming (2008). UK: MIT Press.

[DEG10] De Giusti L, Naiouf M., Chichizola F., Luque E., De Giusti A. "Dynamic Scheduling in Heterogeneous Multiprocessor Architectures. Efficiency Analysis". Computer Science and Technology Series – XV Argentine Congress of Computer Science Selected Papers. La Plata (Buenos Aires): Editorial de la Universidad de La Plata (edulp). 2010. p85 - 95. isbn 978-950-34-0684-7

[DEG16] L. C. De Giusti, F. Chichizola, S. Rodriguez Eguren, M. Sanchez, J. M. Paniego, A. E. De Giusti. "Introduciendo conceptos de Cloud Computing utilizando el entorno CMRE". Proceedings del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016) – Workshop de Innovación en Educación en Informática. Octubre 2016. Pp 1357-1365.

[DEG17] J. Castro, L. D. Giusti, G. Gorga, M. Sánchez, and M. Naiouf. "ECMRE: Extended Concurrent Multi Robot Environment". Computer Science – CACIC 2017. Communications in Computer and Information Science, vol 790, ISBN: 978-3-319-75213-6 978-3-319-75214-3,

- Springer, Cham, págs. 285-294, enero de 2018.
- [DEW15] De Wael, M; Marr, S; De Fraine, B; Van Cutsem, T; De Meuter, W. "Partitioned Global Address Space Languages". *ACM Computing Surveys (CSUR)* 47 (4), 2015.
- [EC213] Amazon Elastic Compute Cloud (Amazon EC2). <http://aws.amazon.com/es/ec2/>. Febrero 2013.
- [GAU15] Adriana Gaudiani. "Simulación y optimización como metodología para mejorar la calidad de la predicción en un entorno de simulación hidrográfica". Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2015.
- [GAU16] A.Gaudiani, E. Luque, P. García, M. Re, M. Naiouf, A. De Giusti. "How a Computational Method Can Help to Improve the Quality of River Flood Prediction by Simulation". *Advances and New Trends in Environmental and Energy Informatics (part V)*. ISBN 978-3-319-23455-7. Pp337-351. 2016.
- [GIL14] Giles MB, Reguly I. 2014 "Trends in high-performance computing for engineering calculations". *Phil.Trans.R.Soc.A* 372: 20130319. <http://dx.doi.org/10.1098/rsta.2013.0319>
- [HAG11] Hager, G. & Wellein, G. *Introduction to HPC for Scientists and Engineers*. (2011) EEUU: CRC Press.
- [JEF13] Jeffers, James; Reinders, James. "Intel Xeon Phi Coprocessor High Performance Programming", Morgan Kaufmann, 2013.
- [KIS13] A. Kishimoto, A. Fukunaga, and A. Botea, "Evaluation of a simple, scalable, parallel best-first search strategy," *Artificial Intelligence*, vol. 195, pp. 222–248, 2013.
- [MCC12] McCool, Michael. "Structured Parallel Programming: Patterns for Efficient Computation", Morgan Kaufmann, 2012
- [MON14] E. Montes de Oca, L. De Giusti, F. Chichizola, A. De Giusti, M. Naiouf. "Utilización de Cluster de GPU en HPC. Un caso de estudio". *Proceedings del XX Congreso Argentino de Ciencias de la Computación (CACIC 2014)*, pp. 1220-1227, 2014.
- [MON16] E. Montes de Oca, L. C. De Giusti, F. Chichizola, A. E. De Giusti, M. Naiouf. "Análisis de uso de un algoritmo de balanceo de carga estático en un Cluster Multi-GPU Heterogéneo". *Proceedings del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016)*, pp. 159-168, 2016.
- [ODR16] Odroid <http://www.hardkernel.com> Accedido 21 de Marzo de 2016.
- [POU16] Pousa A, Sanz V, De Giusti A. "Estructurando código paralelo para clusters heterogéneos de CPUs/GPUs". *Proceedings del XXII Congreso Argentino de Ciencias de la Computación*. ISBN: 978-987-733-072-4. Pp. 139-148
- [RAS16] Raspberry PI. <https://www.raspberrypi.org/> Accedido 21 de Marzo de 2016.
- [RUC15] "Smith-Waterman Protein Search with OpenCL on FPGA". Enzo Rucci, Armando De Giusti, Marcelo Naiouf, Carlos García Sanchez, Guillermo Botella Juan, Manuel Prieto-Matías. *Proceedings of 2014 IEEE Symposium on Parallel and Distributed Processing with Applications (ISPA)*. 20 al 22 de Agosto de 2015. Helsinki, Finlandia. ISBN: 978-1-4673-7952-6. Pp. 208-213. DOI: 10.1109/Trustcom.2015.634.
- [RUC16] "State-of-the-art in Smith-Waterman Protein Database Search". Enzo Rucci, Armando De Giusti, Marcelo Naiouf, Carlos García Sanchez, Guillermo Botella Juan, Manuel Prieto-Matías. *Big Data Analytics in Genomics*. Ka-Chun Wong (Editor). ISBN: 978-3-319-41278-8 (print) 978-3-319-41279-5 (online), Springer, págs. 197-223, 2016.
- [RUC17a] "Accelerating Smith-Waterman Alignment of Long DNA Sequences with OpenCL on FPGA". E. Rucci, C. Garcia, G. Botella, A. De Giusti, M. Naiouf, and M. Prieto-Matias, En: *Bioinformatics and Biomedical Engineering*. IWBBIO 2017. *Lecture Notes in Computer Science*, vol 10209., ISBN: 978-3-319-56154-7, Springer, Cham, págs. 500-511, 2017.
- [RUC17b] "First Experiences Accelerating Smith-Waterman on Intel's Knights Landing Processor". E. Rucci, C. Garcia, G. Botella, A. De Giusti, M. Naiouf, and M. Prieto-Matias, En: *Algorithms and Architectures for Parallel Processing*. ICA3pp 2017. *Lecture Notes in Computer Science*, vol 10393, ISBN: 978-3-319-65482-9, Springer International Publishing, págs. 569-579, 2017.
- [RUC18a] "SWIFOLD:Smith-Waterman Implementation on FPGA with OpenCL for Long DNA Sequences". E. Rucci, C. Garcia, G. Botella, A. De Giusti, M. Naiouf, and M. Prieto-Matias. *BMC Systems Biology*. ISSN: 1752-0509. In press. 2018.
- [RUC18b] "Blocked All-Pairs Shortest Paths Algorithm on Intel Xeon Phi KNL Processor: A Case Study". E. Rucci, A. De Giusti, and M. Naiouf, En: *Computer Science – CACIC 2017. Communications in Computer and Information Science*, vol 790, ISBN: 978-3-319-75213-6 978-3-319-75214-3, Springer, Cham, págs. 47-57, 2018.
- [SAN15] V. Sanz, A. De Giusti, and M. Naiouf, "Performance tuning of the HDA* algorithm for multicore machines." in *Computer Science & Technology Series - XX Argentine Congress of Computer Science - Selected Papers*. Argentina: EDULP, 2015, pp. 51-62.
- [SAN17] Sanz V., De Giusti A., Naiouf M. (2017) "A Hybrid Parallel Search Algorithm for Solving Combinatorial Optimization Problems on Multicore Clusters". *Algorithms and Architectures for Parallel Processing*. ICA3PP 2017. *Lecture Notes in Computer Science*, vol 10393. Springer, Cham.
- [SET13] High-performance Dynamic Programming on FPGAs with OpenCL. Sean Settle. 2013 IEEE High Performance Extreme Computing Conf (HPEC '13), 2013. <https://doi.org/10.1016/j.physa.2017.04.159>
- [BAS17] Basgall, M. J., Hasperué, W., Naiouf, M., & Bariviera, A. F. "Cálculo del exponente de Hurst utilizando Spark Streaming: enfoque experimental sobre un flujo de transacciones de criptomonedas. Presentado en XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017).

Tendencias en Arquitecturas y Algoritmos para Sistemas Paralelos y Distribuidos

Marcelo Naiouf⁽¹⁾, Franco Chichizola⁽¹⁾, Laura De Giusti⁽¹⁾⁽³⁾, Enzo Rucci⁽¹⁾⁽²⁾, Adrián Pousa⁽¹⁾, Ismael Rodríguez⁽¹⁾, Sebastián Rodríguez Eguren⁽¹⁾, Erica Montes de Oca⁽¹⁾, Juan Manuel Paniego⁽¹⁾, Martín Pi Puig⁽¹⁾, Leandro Libutti⁽¹⁾, Javier Balladini⁽⁴⁾, Armando De Giusti⁽¹⁾⁽²⁾

¹Instituto de Investigación en Informática LIDI (III-LIDI),
Facultad de Informática, Universidad Nacional de La Plata – Comisión de Investigaciones Científicas de la
Provincia de Buenos Aires

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CICPBA – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

⁴Universidad Nacional del Comahue

{mnaiouf,francoch,ldgiusti,erucci,apousa,ismael,seguren,emontesdeoca,jmpaniego,mpipuig, libutti, degiusti}@lidi.info.unlp.edu.ar; javier.balladini@gmail.com

Resumen

El eje de esta línea de I/D lo constituye el estudio de tendencias actuales en las áreas de arquitecturas y algoritmos paralelos. Incluye como temas centrales:

- Arquitecturas many-core (GPU, procesadores MIC), FPGAs, híbridas (diferentes combinaciones de multicores y aceleradores), y asimétricas.
- Cloud Computing para HPC (especialmente para aplicaciones de Big Data) y sistemas distribuidos de tiempo real (Cloud Robotics).
- Desarrollo y evaluación de algoritmos paralelos sobre nuevas arquitecturas y su evaluación de rendimiento energético y computacional.

Palabras clave: *Sistemas Paralelos. Clusters. Arquitecturas asimétricas. GPU, MIC, FPGA. Cloud Computing. Cloud robotics. Performance y eficiencia energética.*

Contexto

Se presenta una línea de Investigación que es parte del proyecto del III-LIDI “Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real” (en el marco del Programa de Incentivos del Ministerio de Educación) y de proyectos específicos apoyados por organismos

nacionales e internacionales. También del proyecto “Transformación de algoritmos para nuevas arquitecturas multiprocesador” financiado por la Facultad de Informática de la UNLP.

En el tema hay cooperación con varias Universidades de Argentina y se está trabajando con Universidades de América Latina y Europa en proyectos financiados por CyTED, AECID y la OEI (Organización de Estados Iberoamericanos).

Por otra parte, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado y se ha tenido el apoyo de diferentes empresas (IBM, Microsoft, Telecom, Intel) en la temática de Cloud Computing.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Asimismo, el III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD) del MinCyT.

Introducción

Una de las áreas de creciente interés lo constituye el cómputo de altas prestaciones, en el cual el rendimiento está relacionado con dos aspectos: por un lado las arquitecturas de soporte y por el otro los algoritmos que hacen uso de las mismas.

A la aparición de arquitecturas *many-core* (como las GPU o los procesadores MIC), se

ha sumado el uso de FPGAs debido a su potencia de cómputo y rendimiento energético. Su combinación en sistemas HPC da lugar a plataformas híbridas con diferentes características [RUC16].

Lógicamente, esto trae aparejado una revisión de los conceptos del diseño de algoritmos paralelos (incluyendo los lenguajes mismos de programación y el software de base), así como la evaluación de las soluciones que éstos implementan. También resulta necesario investigar las estrategias de distribución de datos y procesos a fin de optimizar la performance.

Además, el estudio del consumo y la eficiencia energética de los nuevos sistemas paralelos se vuelve tan importante como el de métricas clásicas (speedup, eficiencia, escalabilidad) debido a los costos económicos y los problemas operativos asociados [BAL13].

Por otra parte, los avances en las tecnologías de virtualización y cómputo distribuido han dado origen al paradigma de Cloud Computing, que se presenta como una alternativa a los tradicionales sistemas de Clusters y Multicluster para ambientes de HPC [BER08]. A su vez, este concepto se puede ampliar a sistemas distribuidos de tiempo real, en particular sistemas inteligentes como son los robots que pueden trabajar en paralelo utilizando su propia capacidad de procesamiento y al mismo tiempo conectándose con la potencia de un servidor en la nube (*Cloud Robotics*) [LOR13][GUO12][KEH15].

En esta línea de I/D se trabaja sobre aspectos que marcan tendencias en el área.

GPUs y Cluster de GPUs

Las GPUs son el acelerador dominante en la comunidad de HPC al día de hoy por su alto rendimiento y bajo costo de adquisición. En la actualidad, tanto NVIDIA como AMD trabajan especialmente en mejorar la eficiencia energética de sus placas y disminuir el alto costo de programación.

La combinación de GPUs con otras plataformas paralelas como clusters y multicores, brindan un vasto conjunto de

posibilidades de investigación en arquitecturas híbridas, a partir de diferentes combinaciones a saber:

- Máquinas multicore con más de una GPU, que combinan herramientas de programación paralela como OpenMP/CUDA o Pthread/CUDA.
- Cluster de máquinas multicore cada una con una o más placas de GPU, lo que permite combinar OpenMP/MPI/CUDA o Pthread/MPI/CUDA.

Los desafíos que se plantean son múltiples, sobre todo en lo referido a distribución de datos y procesos en tales arquitecturas híbridas a fin de optimizar el rendimiento de las soluciones.

MIC

En forma reciente Intel brinda una alternativa a partir de la arquitectura MIC (*Many Integrated Core Architecture*). Esta arquitectura permite utilizar métodos y herramientas estándar de programación con altas prestaciones (lo que los distingue especialmente de las GPUs). De esta forma, se remueven barreras de entrenamiento y se permite focalizar en el problema más que en la ingeniería del software. Xeon Phi es el nombre elegido por Intel para su serie de procesadores many-core. Recientemente, Intel ha lanzado Knights Landing (KNL), la segunda generación de Xeon Phi. A diferencia de sus predecesores que operaban como co-procesador a través del puerto PCI, los procesadores KNL pueden operar en forma autónoma. Además, integran las nuevas extensiones vectoriales AVX-512 y tecnología de memoria 3D, entre otras características avanzadas [REI16].

FPGAs

Una FPGA (*Field Programmable Gate Array*) es una clase de acelerador basado en circuitos integrados reconfigurables. La capacidad de adaptar sus instrucciones de acuerdo a la aplicación objetivo le permite incrementar la productividad de un sistema y mejorar el rendimiento energético para ciertos tipos de aplicaciones. Tradicionalmente fueron utilizadas para el procesamiento digital de

señales. Sin embargo, en los últimos años, existen dos tendencias claras para extender su uso a otros dominios. En primer lugar, el establecimiento de alianzas estratégicas entre fabricantes de procesadores y de FPGAs para integrar estos dispositivos en arquitecturas híbridas (Intel con Altera; IBM con Xilinx) [IBM15][INT16]. En segundo lugar, el desarrollo de nuevas herramientas de programación para FPGAs empleando estándares familiares para HPC, con las cuales se espera reducir los tradicionales tiempos y costos de programación [SEA13][XIL15]. Por último, la incorporación de FPGAs a los servicios de cloud abre nuevas oportunidades para la explotación de esta clase de aceleradores.

Eficiencia energética

La mejora de la eficiencia energética es una de las principales preocupaciones en la informática actual, principalmente a partir de las plataformas con gran cantidad de procesadores. Muchos esfuerzos están orientados a tratar la eficiencia energética como eje de I/D, como métrica de evaluación, y también a la necesidad de metodologías para medirla.

Entre los puntos de interés pueden mencionarse:

- Análisis de metodologías y herramientas para medir consumo energético.
- Estudio de técnicas para reducir el consumo energético en aplicaciones de HPC de acuerdo a las arquitecturas utilizadas.
- Evaluación de eficiencia energética de diferentes algoritmos y plataformas paralelas.
- Optimización de la eficiencia energética. A partir de los valores de energía que brindan los contadores hardware es posible definir estrategias de programación que lleven a reducir el consumo, manteniendo a su vez el rendimiento en valores aceptables [SAE17].

Cloud Computing

Cloud Computing, proporciona grandes conjuntos de recursos físicos y lógicos (como pueden ser infraestructura, plataformas de desarrollo, almacenamiento y/o aplicaciones), fácilmente accesibles y utilizables por medio

de una interfaz de administración web, con un modelo de arquitectura “virtualizada” [SHA10][XIN12]. Estos recursos son proporcionados como servicios (“as a service”) y pueden ser dinámicamente reconfigurados para adaptarse a una carga de trabajo variable (escalabilidad), logrando una mejor utilización y evitando el sobre o sub dimensionamiento (elasticidad) [VEL09].

Más allá de las potenciales características y beneficios que brinda un Cloud, de por sí atractivas, es de gran interés estudiar el despliegue de entornos de ejecución para cómputo paralelo y distribuido (Clusters Virtuales), como así también realizar I/D en la portabilidad de las aplicaciones de HPC en el Cloud [DOE11][ROD11].

Por otro lado, Cloud Robotics es una de las áreas más prometedoras de la investigación informática actual en la cual se cuenta con “robots” dotados de diferentes sensores y capacidades, conectados a un Cloud vía Internet. Los temas de investigación derivados son múltiples: sensores, redes de sensores e inteligencia distribuida; robótica y sistemas colaborativos de tiempo real basados en robots; aplicaciones críticas (por ej. en ciudades inteligentes o en el ámbito industrial).

Dispositivos de bajo costo con capacidades para cómputo paralelo

En la actualidad se comercializan placas de bajo costo como Raspberry PI [RAS17] u Odroid [ODR17] que poseen múltiples núcleos simples. Asimismo, existen diversos dispositivos móviles con capacidades similares. Es de interés estudiar como explotar el paralelismo en estos dispositivos para mejorar el rendimiento y/o consumo energético de las aplicaciones.

Líneas de Investigación, Desarrollo e Innovación

- Arquitecturas many-core (procesadores MIC y GPU) y FPGA. Análisis de este tipo de máquinas y de técnicas para desarrollar código optimizado.

- Arquitecturas híbridas (diferentes combinaciones de clusters, multicores, manycores y FPGAs). Diseño de algoritmos paralelos sobre las mismas.
- Exploración de nuevos lenguajes y modelos de programación para HPC.
- Cloud Computing para realizar HPC. Evaluación de performance en este tipo de arquitectura. Análisis del overhead por el software de administración del Cloud.
- Sistemas inteligentes de tiempo real distribuidos (Cloud Robotics).
- Interconexión de Brokers de mensajes MQTT sobre Cloud Públicos y Cloud Privados.
- Consumo energético en las diferentes arquitecturas paralelas, en particular en relación a los algoritmos paralelos y la configuración de la arquitectura. Análisis de metodologías y herramientas de medición.

Resultados y Objetivos

Investigación experimental a realizar

- Desarrollo y evaluación de algoritmos paralelos sobre nuevas arquitecturas. Análisis de rendimiento, eficiencia energética y costo de programación.
- Análisis del overhead introducido por el sistema gestor del Cloud en un entorno de HPC para aplicaciones científicas de Big Data. Comparar el rendimiento entre Cloud y Cluster Computing.
- Analizar y comparar las diferentes estrategias para interconectar brokers de mensajes MQTT tanto sobre cloud públicos como privados.
- Realizar el desarrollo de nuevos planificadores de tareas para multicores asimétricos sobre diferentes sistemas operativos con el objetivo de maximizar el rendimiento y minimizar el consumo de energía [SAE15][SAE17].
- Optimización de algoritmos paralelos para controlar el comportamiento de múltiples robots que trabajan colaborativamente, considerando la distribución de su capacidad de procesamiento “local” y la coordinación con la potencia de cómputo y capacidad de

almacenamiento (datos y conocimiento) de un Cloud.

Resultados obtenidos

- Se ha finalizado una tesis de doctorado en la línea de planificación de procesos sobre multicores asimétricos (AMPs). Se analizó la relación entre rendimiento, justicia y eficiencia energética al ejecutar varias cargas de trabajo diversas sobre AMPs. El análisis permitió el diseño de nuevas estrategias de planificación, implementadas sobre sistemas operativos de propósito general, orientadas a optimizar estos objetivos [POU17].
- Se estudiaron técnicas de programación y optimización para la nueva generación de procesadores Xeon Phi (*Knights Landing*) [RUC18].
- Se cuantificaron las mejoras introducidas por el uso de GPUs para aplicaciones de procesamiento digital de imágenes tanto desde el punto de vista del desempeño como del consumo energético [PIP17].
- Se evaluó el uso de diferentes arquitecturas paralelas (CPU, Xeon Phi, GPU, FPGA) para procesamiento de grandes conjuntos de datos biológicos considerando no sólo el rendimiento sino la eficiencia energética [RUC17a][RUC17b].
- Se analizó la precisión en la predicción del consumo de energía por parte de la herramienta Intel RAPL [PAN18].
- Análisis del despliegue de un sistema multi-robot, integrado por un chasis de auto Rover de 4 ruedas y un cuadricóptero Parrot Bebop, conectados a un servicio de cloud público [DEG17].

Organización de Eventos

En el año 2017 se han organizado las V Jornadas de Cloud Computing y Big Data (JCC&BD 2017) en Argentina, con participación de especialistas académicos del país y del exterior y de empresas con experiencia en Cloud Computing. En junio de 2018 se organizarán las VI Jornadas de Cloud Computing y Big Data (JCC&BD 2018).

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D el último año se concluyeron: 1 tesis doctoral, 1 tesis de maestría y 1 trabajo de Especialización. Al mismo tiempo se encuentran en curso 2 tesis de Doctorado en Ciencias Informáticas y 2 tesis de Maestría.

Además, se participa en el dictado de las carreras de Doctorado en Ciencias Informáticas, y Magíster y Especialización en Cómputo de Altas Prestaciones de la Facultad de Informática de la UNLP (acreditadas por la CONEAU con categoría A, B y A, respectivamente), por lo que potencialmente pueden generarse nuevas Tesis de Doctorado y Maestría y Trabajos Finales de Especialización.

Existe cooperación con grupos de otras Universidades del país y del exterior, y hay tesistas de diferentes Universidades realizando su Tesis con el equipo del proyecto.

Respecto a las carreras de grado, se dictan por parte de integrantes de la línea de investigación tres materias directamente relacionadas con los temas de la misma: “Taller de Programación sobre GPUs”, “Cloud Computing y Cloud Robotics” y “Conceptos y Aplicaciones en Big Data”.

Referencias

[BAL13] Ballardini, J., Rucci, E., De Giusti, A., Naiouf, M., Suppi, R., Rexachs, D., Luque, E. “Power characterisation of shared-memory HPC systems”. *Computer Science & Technology Series – XVIII Argentine Congress of Computer Science Selected Papers*. ISBN 978-987-1985-20-3. Págs. 53-65. 2013.

[BER08] Bertogna, M., Grosclaude, E., Naiouf, M., De Giusti, A., Luque, E.: “Dynamic on Demand Virtual Clusters in Grids”. In: 3rd Workshop on Virtualization in High-Performance Cluster and Grid Computing (VHPC 08). España. (2008).

[BOR05] S. Y. Borkar, P. Dubey, K. C. Kahn, D. J. Kuck, H. Mulder, S. S. Pawlowski, y J. R. Rattner: “Platform 2015: Intel® Processor and Platform Evolution for the Next Decade”. In: Intel Corporation, White Paper, 2005.

[DEG17] A. E. De Giusti, I. P. Rodriguez, M. Costanzo, and M. Boggia. “Cloud robotics: Auto Rover 4wd y cuadricóptero controlados remotamente desde AWS”. *Actas del XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017)*., ISBN: 978-950-34-1539-9, págs. 165-174, octubre de 2017.

[DOE11] Doelitzcher, F., Held, M., Sulistio, A., Reich, C. ViteraaS: “Virtual Cluster as a Service”. In: 3rd IEEE International Conference on Cloud Computing Technology and Science. Grecia (2011).

[GUO12] Guoqiang Hu, Wee Peng Tay, Yonggang Wen: “Cloud robotics: architecture, challenges and applications”. In: *Network, IEEE*, vol.26, no.3, pp.21-28. 2012.

[IBM15] IBM. “IBM and Xilinx Announce Strategic Collaboration to Accelerate Data Center Applications”. Disponible en <https://www-03.ibm.com/press/us/en/pressrelease/48074.wss>

[INT16] Intel. “Intel Acquisition of Altera”. Disponible en <http://intelacquiresaltera.transactionannouncement.com>

[KEH15] Kehoe B., Patil S., Abbeel P., Goldberg K.: “A Survey of Research on Cloud Robotics and Automation”. In: *IEEE Transactions on Automation Science and Engineering (T-ASE): Special Issue on Cloud Robotics and Automation*. Vol. 12, no. 2. 2015.

[LOR13] Lorencik D., Sincak P.: “Cloud robotics: Current trends and possible use as a service”. In: *Applied Machine Intelligence and Informatics (SAMII)*, 2013 IEEE 11th International Symposium on , vol., no., pp.85-88. 2013.

[ODR16] Odroid <http://www.hardkernel.com> Accedido 21 de Marzo de 2016.

[PAN18] J. M. Paniego, S. Gallo, M. P. Puig, F. Chichizola, L. D. Giusti, and J. Ballardini. “Analysis of RAPL Energy Prediction Accuracy in a Matrix Multiplication Application on Shared Memory”. En: *Computer Science – CACIC 2017. Communications in Computer and Information Science*, vol 790, ISBN: 978-3-319-75213-6 978-3-319-75214-3, Springer, Cham, págs. 37-46, enero de 2018.

[PIP17] M. Pi Puig, L. C. De Giusti, M. Naiouf, and A. E. De Giusti. “GPU Performance and Power Consumption Analysis: A DCT based denoising application” *Actas del XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017)*., ISBN: 978-950-34-1539-9, págs. 185-195, octubre de 2017.

[POU17] Pousa Adrian, De Giusti Armando, Saez Alcaide Juan Carlos. Tesis de Doctorado: “Optimización de rendimiento, justicia y consumo energético en sistemas multicore asimétricos mediante planificación” <http://sedici.unlp.edu.ar/handle/10915/62960>

[RAS16] Raspberry PI. <https://www.raspberrypi.org/> Accedido 21 de Marzo de 2016.

[REI16] Reinders, J., Jeffers, J., Sodani, A. “Intel Xeon Phi Processor High Performance Programming Knights Landing Edition”. Morgan Kaufmann Publishers Inc., Boston, MA, USA, 2016

[ROD11] Rodriguez, I., Pettoruti, J.E., Chichizola, F., De Giusti, A.: “Despliegue de un Cloud Privado para entornos de cómputo científico”. In: *Proceedings del XVII Congreso Argentino de Ciencias de la Computación*. Argentina (2011).

[RUC16] Rucci, Enzo: “Evaluación de rendimiento y eficiencia energética en sistemas heterogéneos para

- bioinformática”. Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016.
- [RUC17a] E. Rucci, C. Garcia, G. Botella, A. De Giusti, M. Naiouf, and M. Prieto-Matias. “Accelerating Smith-Waterman Alignment of Long DNA Sequences with OpenCL on FPGA”. En: Bioinformatics and Biomedical Engineering. IWBBIO 2017. Lecture Notes in Computer Science, vol 10209., ISBN: 978-3-319-56154-7, Springer, Cham, págs. 500-511, 2017.
- [RUC17b] E. Rucci, C. Garcia, G. Botella, A. De Giusti, M. Naiouf, and M. Prieto-Matias. “First Experiences Accelerating Smith-Waterman on Intel’s Knights Landing Processor”. En: Algorithms and Architectures for Parallel Processing. ICA3pp 2017. Lecture Notes in Computer Science, vol 10393, ISBN: 978-3-319-65482-9, Springer International Publishing, págs. 569-579, 2017.
- [RUC18] “Blocked All-Pairs Shortest Paths Algorithm on Intel Xeon Phi KNL Processor: A Case Study”. E. Rucci, A. De Giusti, and M. Naiouf, En: Computer Science – CACIC 2017. Communications in Computer and Information Science, vol 790, ISBN: 978-3-319-75213-6 978-3-319-75214-3, Springer, Cham, págs. 47-57, 2018.
- [SAE15] Juan Carlos Saez, Adrian Pousa, Daniel Chaver, Fernando Castro, Manuel Prieto Matias: “ACFS: A Completely Fair Scheduler for Asymmetric Single-ISA Multicore Systems”. In: ACM SAC 2015 (The 30TH ACM/SIGAPP Symposium on applied computing). 2015.
- [SAE17] Saez, J.C., Pousa, A., Rodríguez-Rodríguez, R., Castro, F., Prieto-Matias, M. “PMCTrack: Delivering performance monitoring counter support to the OS scheduler”. The computer journal Volume 60, Issue 1 January 2017.
- [SEA13] Sean Settle: “High-performance Dynamic Programming on FPGAs with OpenCL”. In: IEEE High Performance Extreme Computing Conference. 2013.
- [SHA10] Shafer, J.: “I/O virtualization bottlenecks in cloud computing today”. In: Proceedings of the 2nd conference on I/O virtualization (VIOV10). USA (2010).
- [VEL09] Velte, A.T., Velte, T.J., Elsenpeter, R.: “Cloud Computing: A Practical Approach”. McGraw Hill Professional. 2009.
- [XIL15] Xilinx Inc. “SDAccel Development Environment”. [Online]. Disponible en <http://www.xilinx.com/products/design-tools/software-zone/sdaccel.html>
- [XIN12] Xing, Y., Zhan, Y.: “Virtualization and Cloud Computing”. In: Proceedings pp.305-312, Springer Link. ISBN 978-3-642-27323-0. (2012). Morgan Kaufmann. 2013.

Aplicaciones de Cómputo Intensivo con Impacto Social

Javier Balladini¹, Marina Morán¹, Pablo Bruno¹, Belén Casanova¹, Claudia Rozas¹,
Federico Uribe^{1,2}, Sebastián Gomez², Nestor Vicente³, Cristina Orlandi³, Armando De Giusti⁴,
Remo Suppi⁵, Dolores Rexachs⁵, Emilio Luque⁵

¹ Facultad de Informática, Universidad Nacional del Comahue

{javier.balladini, marina, claudia.rozas, federico.uribe}@fi.uncoma.edu.ar; psbruno90@gmail.com;
mb.casanova.retamal@gmail.com

² Poder Judicial del Neuquén

sebastian.gomez@jusneuquen.gov.ar

³ Hospital Francisco Lopez Lima

nestorvicente07@yahoo.com.ar; orlandi.mariacristina@gmail.com

⁴ Instituto de Investigación en Informática LIDI, Universidad Nacional de La Plata

degiusti@lidi.info.unlp.edu.ar

⁵ Departamento de Arquitectura de Computadores y Sistemas Operativos, Universidad Autónoma de Barcelona

{remo.suppi, dolores.rexachs, emilio.luque}@uab.es

Resumen

Los avances tecnológicos de los sistemas de cómputo paralelo y distribuido permiten el desarrollo de aplicaciones antes impensadas. Nuestra investigación se centra en desarrollar metodologías, modelos y soluciones informáticas para colaborar en la resolución de problemas que tengan una alta demanda computacional e impacto social. Hemos definido tres ejes de investigación: aplicaciones para la salud, aplicaciones de informática forense, y consumo energético de los sistemas de cómputo paralelo. Estas líneas de investigación se desarrollan en colaboración con una universidad nacionales y otra del extranjero, un hospital público, y un gabinete provincial de informática forense.

Palabras claves: computación de altas prestaciones, eficiencia energética, aplicaciones para la salud, aplicaciones de informática forense.

1. Contexto

La línea de investigación aquí presentada está enmarcada dentro del proyecto de investigación 04/F013 "Aplicaciones de Cómputo Intensivo con Impacto Social", financiado por la Universidad Nacional del Comahue (UNComa), con inicio el 01/01/2017 y finalización el 31/12/2020, y acreditado por el Ministerio de Educación de Argentina. Otra fuente de financiamientos es el Inter-U "Colaboración UNComa-UNLP: docencia e investigación en Sistemas Paralelos" de PROMINF.

Uno de los tres ejes centrales de nuestra investigación, las aplicaciones para la salud, se desarrolla en colaboración con la Unidad de Terapia Intensiva y la Unidad de Vigilancia Intermedia, ambos pertenecientes al "Hospital Francisco López Lima" de la ciudad de General Roca, provincia de Río Negro. El eje de aplicaciones de informática forense, se desarrolla dentro del marco de colaboración con el Gabinete de Pericias Informáticas del Poder Judicial de la provincia de Neuquén. Respecto al eje relacionado con el consumo energético de los sistemas de cómputo paralelo, se desarrolla en colaboración con el Instituto de Investigación en Informática LIDI de la Universidad Nacional de La Plata (UNLP), y el grupo de investigación HPC4EAS (High Performance Computing for Efficient Applications and Simulation) de la Universidad Autónoma de Barcelona (UAB) de España.

2. Introducción

Una de las áreas de mayor interés en la actualidad es la Computación de Altas Prestaciones (HPC, del inglés, *High Performance Computing*). La computación paralela es un tipo de computación en el que los cálculos se realizan de forma simultánea. Si bien el paralelismo ha sido empleado históricamente en la computación de altas prestaciones, ha ganado un enorme interés debido al impedimento para seguir aumentando la frecuencia de reloj de los procesadores; el problema se encuentra en el alto consumo energético y disipación del calor a altas frecuencias. Como

no se podía seguir aumentando la frecuencia para que las aplicaciones ejecuten más rápido, la solución fue incrementar la cantidad de unidades de procesamiento, dando así lugar a la aparición de procesadores multinúcleos. Desde entonces, la computación paralela se ha convertido en el paradigma dominante en la arquitectura de computadoras.

Para algunas aplicaciones, será suficiente con utilizar una plataforma comprendida por una única computadora con uno o más procesadores multinúcleos. En otros casos, podrá ser necesario el poder de cómputo de una agregación de computadoras, como por ejemplo del tipo *cluster*. La masificación de la tecnologías de cómputo paralelo hacen que ellas sean cada vez más accesibles, y se pueda pensar en el desarrollo de nuevas aplicaciones, muchas de las cuales pueden tener un fuerte impacto para el beneficio social. Sin embargo, extraer el máximo rendimiento de estas plataformas requiere utilizar técnicas específicas de programación paralela, que son más difíciles que las típicas de programación secuencial, principalmente debido a la sincronización, comunicación de tareas, y complejidad de la arquitectura de las plataformas hardware.

Nuestro interés está centrado en tres áreas temáticas: aplicaciones de cómputo intensivo para la salud, aplicaciones de cómputo intensivo de informática forense, y el consumo energético de los sistemas de HPC. A continuación se introduce cada uno de estos temas:

Aplicaciones para la salud

En los centros de salud, la Unidad de Cuidados Intensivos (UCI) atiende a pacientes cuya salud está en condiciones críticas, con riesgo de muerte. Por esta razón deben contar con asistencia médica las 24 horas del día y ser controlados en forma rigurosa; mínimamente están conectados a un monitor que mide sus signos vitales y da alertas, considerando aquellos aspectos que pueden indicar un riesgo para la salud del paciente bajo criterios aplicados según la población estándar.

El equipamiento utilizado para monitorear a los pacientes y realizar diferentes estudios ha sido desarrollado para trabajar en forma independiente, sin contemplar la posibilidad de incorporar información adicional obtenida a través de otros medios (ya sea otros estudios, datos inherentes del paciente, etc). Es decir, no se ha previsto un uso integral de toda la información del paciente, sino que los estudios se realizan en forma aislada y luego el personal médico debe realizar el análisis de los mismos considerando todas las variables conocidas. Esta metodología

demora la detección de patologías. Además, debido a que el equipamiento médico generalmente no está preparado para el registro histórico de datos, el personal de la UCI toma registros manuales en intervalos de horas. Así, la mayoría de las mediciones que realiza el equipamiento médico se pierden, y esta omisión de información podría reducir la precisión del diagnóstico.

Un gran avance sería disponer de un sistema de cómputo que detecte patologías en tiempo real, basándose en múltiples parámetros de diferentes medios. La detección temprana de patologías permitirá aumentar la efectividad de los tratamientos, y por consiguiente la mejora de la salud de los pacientes y la reducción del costo económico. El problema principal que debe enfrentarse para la construcción de este sistema, y que creemos viable con la aplicación de técnicas de computación paralela, es el procesamiento en tiempo real de un gran volumen de datos generado por el equipamiento (especialmente las curvas como, por ejemplo, el electrocardiograma).

No hay muchos sistemas de este tipo, algunos de ellos se encuentran en etapa experimental inicial y otros ya llevan algunos años de investigación. La información disponible de estos sistemas es normalmente escasa por tratarse mayormente de software privativo. En la bibliografía se encuentran algunos trabajos como [13, 3, 1, 2].

Aplicaciones de informática forense

En la actualidad, una gran parte de las investigaciones judiciales involucran elementos de prueba que son potenciales fuentes de evidencia digital [11, 10]. El volumen de datos digital que debe ser sometido a análisis forense está aumentando de forma sostenida, incrementando el tiempo de procesamiento requerido para el análisis. Este aumento en el tiempo de análisis forense lleva a un crecimiento excesivo de la lista de espera de pericias en trámite e impacta negativamente en la investigación que deben llevar a cabo los organismos jurisdiccionales.

Muchas tareas periciales involucran un elevado tiempo de procesamiento: detección de imágenes sospechosas de contener pornografía, la generación de índices para realizar consultas dinámicas sobre el corpus digital, la creación de diccionarios personalizados para descifrado de contraseñas en base a la información digital contenida en los dispositivos de almacenamiento, la generación de listas de valores hash para análisis de firmas, la detección de malware o localización de archivos relevantes a la investigación, y la búsqueda de evidencia digital mediante palabras clave. En especial, la detección de imágenes con contenido pornográfico [8] es una de las tareas

que requieren mayor tiempo de procesamiento.

Las herramientas disponibles en el mercado no satisfacen los tiempos de procesamiento requeridos para estas actividades. Los avances tecnológicos y técnicas de programación de los sistemas de cómputo paralelo pueden hacer viable el desarrollo de aplicaciones para informática forense que tengan tiempos de procesamiento menores a los actuales. En muchos casos, las herramientas también carecen de capacidades de ejecución no interactiva que posibilite la automatización de tareas para ser utilizadas en entornos de cómputo paralelo y distribuido.

Consumo energético

Mientras el rendimiento de los sistemas de computación de altas prestaciones (HPC, High Performance Computing) continúa creciendo, las máquinas aumentan significativamente la cantidad de unidades de procesamiento. Este aumento en el número de componentes hace disminuir la confiabilidad y aumentar el consumo energético de un sistema de cómputo. Así, a pocos años de arribar a la era exaescala (prevista para 2020), el consumo energético se han identificado como uno de los mayores desafíos a enfrentar [14, 12].

El consumo energético es hoy en día un gran problema. Para dar una idea de la magnitud del mismo, utilizaremos de ejemplo la máquina de mayor prestaciones de la actualidad, la máquina China Sunway TaihuLight. Esta máquina demanda 15 MW de potencia, lo mismo que se requiere para abastecer a los hogares de una ciudad con alrededor de 200.000 habitantes (cálculo realizado en base al consumo de un hogar en Argentina). Además del alto impacto económico, la generación de tanta energía podría tener un alto impacto medioambiental, por ejemplo, represas hidroeléctricas que modifican el ecosistema, y social, por ejemplo, la mayor fuente de energía mundial se obtiene del carbón, cuya extracción minera es altamente peligrosa.

La computación ecológica es el estudio y la práctica de la computación ambientalmente sostenible. Ella se ocupa de diferentes aspectos de los sistemas de cómputo: diseño, manufactura, eliminación, y uso. Este último aspecto, el uso ecológico, se refiere al uso de los sistemas de cómputo con conciencia ambiental. Es posible reducir el consumo de energía de los sistemas de cómputo utilizando diferentes estrategias que deben ser consideradas a nivel del software, y consisten en realizar cambios en la configuración del sistema o en las aplicaciones. Estas estrategias incluyen: explotación del paralelismo (muchos cores lentos consume menos energía que po-

cos cores rápidos), uso adecuado de la jerarquía de memoria, hibernación de recursos, escalado dinámico de frecuencia y tensión, rediseño de algoritmos, planificación de tareas, asignación de tareas a recursos hardware.

3. Líneas de investigación

El eje central de nuestra investigación es desarrollar metodologías, modelos y soluciones informáticas para colaborar en la resolución de problemas que tengan una alta demanda computacional e impacto social en los siguientes campos: aplicaciones para la salud, aplicaciones de informática forense, consumo energético de los sistemas de cómputo paralelo.

Aplicaciones para la salud

Esta línea está enfocada en el estudio y desarrollo de un sistema para detección temprana de patologías en Unidades de Terapia Intensiva (UTI), y que puede también abarcar a las Unidades de Vigilancia Intermedia (UVI). Nuestro objetivo está orientado a resolver el procesamiento de una gran cantidad de datos en tiempo real, proveniente de señales de equipamiento médico, utilizando técnicas de computación paralela.

Aplicaciones de informática forense

Esta línea se centra en acelerar el procesamiento de grandes volúmenes de datos, principalmente de videos e imágenes, mediante técnicas de computación paralela. En especial, interesa acelerar el proceso de detección automática de imágenes con pornografía.

Consumo energético

Nos centramos en el desarrollo de metodologías, modelos y construcción de software para administrar y gestionar el consumo de energía y prestaciones de sistemas de cómputo paralelo. Nuestros objetivos principales son:

- Predicción de energía y rendimiento. Es importante proveer a un administrador de sistema de herramientas que permitan predecir la energía y el rendimiento que producirían distintas configuraciones del sistema al ejecutar una dada aplicación paralela, y así poder seleccionar la configuración adecuada que mantenga el compromiso deseado entre tiempo de ejecución y eficiencia energética.
- Gestión energética en mecanismos de tolerancia a fallos. La tolerancia a fallos agrega

una carga de trabajo significativa al sistema de cómputo, sobre todo en sistemas que tienen enormes cantidades de unidades de procesamiento [9], haciendo necesario gestionar el consumo energético de los distintos mecanismos.

4. Resultados y objetivos

Aplicaciones para la salud

Los objetivos específicos en curso son:

- Desarrollar hardware y software para extraer datos del equipamiento médico, y transmitirlos por WiFi a la plataforma de procesamiento de la información. La extracción de datos no es trivial debido al uso de protocolos de comunicaciones propietarios que los fabricantes no dan a conocer.
- Desarrollar aplicaciones para el procesamiento eficiente de señales (como el electrocardiograma). Éstas deben ejecutar en máquinas con procesadores de propósito general (CPUs), y utilizar los recursos de cómputo de manera eficiente para reducir el tamaño de la plataforma hardware que requiere el sistema. Se incluye también la detección de anomalías en las señales para evitar la contaminación del sistema con datos erróneos.
- Diseñar la infraestructura de un sistema de Big Data que permita el almacenamiento, análisis y procesamiento en tiempo real de señales extraídas del equipamiento médico (monitores de signos vitales, respiradores, etc.) y otros datos ingresados manualmente.
- Desarrollar una aplicación para la interacción con médicos y enfermeros.

Los avances/resultados actualmente comprenden:

- Un análisis del estado general de las UTI del hospital Francisco Lopez Lima [7].
- El desarrollo de un dispositivo embebido que obtiene información de la señal del electrocardiograma de un monitor médico, a través de una salida analógica, y la transmite por WiFi para su posterior procesamiento.
- El desarrollo parcial de una aplicación que detecta complejos QRS de manera eficiente en señales de electrocardiogramas.
- Análisis parciales de rendimientos de diferentes alternativas tecnológicas de la infraestructura de Big Data.

Aplicaciones de informática forense

Se han evaluado distintas alternativas de software para detección de pornografía disponible en la literatura, y se ha seleccionado la aplicación desarrollada por Yahoo [4] debido a que posee una alta tasa de aciertos. Actualmente, se está trabajando en el análisis de la aplicación seleccionada, y en su optimización para acelerar el tiempo de procesamiento de imágenes en plataformas paralelas basadas en CPUs. Además, se espera avanzar en la ejecución distribuida de la aplicación de detección de imágenes pornográficas para aumentar la productividad en el procesamiento de una cantidad masiva de imágenes, y en el desarrollo de una aplicación que permita el lanzamiento de la aplicación para el procesamiento de un conjunto específico de imágenes y generación de reportes forenses.

Consumo energético

Tras dar los primeros pasos en la predicción de energía y rendimiento para aplicaciones SPMD construidas con el modelo de programación de paso de mensajes (MPI) [6], actualmente se está avanzando en la predicción energética para aplicaciones SPMD implementadas con un modelo de programación híbrido de paso de mensajes (MPI) y memoria compartida (OpenMP), y en una metodología de predicción mejorada que permita una mayor precisión.

Los métodos de tolerancia a fallos tienen fuerte incidencia en el consumo energético de los sistemas de HPC, y resulta de suma importancia conocer, antes de ejecutar una cierta aplicación, el impacto que pueden producir los diferentes métodos y configuraciones del mismo. En [5], presentamos una metodología para predecir el consumo energético producido por el método de checkpoint coordinado remoto. Actualmente, hemos agregado la predicción para la operación de restart (adicionalmente a la de checkpoint), y se están contemplando distintas alternativas de configuraciones del sistema, relacionadas a: almacenamiento en NFS, configuración de la aplicación de checkpoint/restart, y energéticas (estados C y P de las CPUs). A futuro, esperamos extender el trabajo a la propuesta de mecanismos de gestión de tolerancia a fallos que procuren un uso eficiente del cluster, permitiendo maximizar la productividad y minimizar el consumo energético.

5. Formación de recursos humanos

El equipo de trabajo cuenta con un integrante doctorado en la temática (año 2008). En 2018 se espera la finalización de dos tesis de grado en temas de aplicaciones para la salud, y en 2019 una tesis doctoral en el tema de consumo energético.

Referencias

- [1] Amara health analytics, <http://www.amarahealthanalytics.com> (accedido en marzo de 2018).
- [2] ehcos smarticu, <http://www.ehcos.com/productos/ehcos-smarticu/> (accedido en marzo de 2018).
- [3] Excel medical, <http://excel-medical.com/> (accedido en marzo de 2018).
- [4] Open sourcing a deep learning solution for detecting nsfw images, <https://yahooeng.tumblr.com/post/151148689421/open-sourcing-a-deep-learning-solution-for> (accedido en marzo de 2018).
- [5] Javier Balladini, Marina Morán, Dolores Rexachs, and Emilio Luque. Metodología para predecir el consumo energético de checkpoints en sistemas de hpc. *XX Congreso Argentino de Ciencias de la Computación (CACIC 2014)*, 2014.
- [6] Javier Balladini, Ronal Muresano, Remo Suppi, Dolores Rexachs, and Emilio Luque. Methodology for predicting the energy consumption of spmd application on virtualized environments. *Computer Science and Technology (JCST)*, 13(3):130–136, 2013.
- [7] Javier Balladini, Claudia Rozas, Emmanuel Frati, Nestor Vicente, and Cristina Orlandi. Big data analytics in intensive care units: challenges and applicability in an argentinian hospital. *Computer Science and Technology (JCST)*, 2015.
- [8] Carlos Caetano, Sandra Eliza Fontes de Avila, William Robson Schwartz, Silvio Jamil Ferzoli Guimarães, and Arnaldo de Albuquerque Araújo. A mid-level video representation based on binary descriptors: A case study for pornography detection. *Neurocomputing*, 213:102–114, 2016.
- [9] Franck Cappello, Al Geist, William Gropp, Sanjay Kale, Bill Kramer, and Marc Snir. Toward exascale resilience: 2014 update. *Supercomputing Frontiers and Innovations*, 1(1), 2014.
- [10] Simson L. Garfinkel. Digital forensics research: The next 10 years. *Digit. Investig.*, 7:S64–S73, August 2010.
- [11] Sebastián Gómez, Hernán Herrera, and Federico Uribe. Automatización y computación distribuida para laboratorios de informática forense. In *45^o JAIIO - Jornadas Argentinas de Informática*, 2016, CABA, Argentina.
- [12] Robert Lucas, James Ang, Keren Bergman, Shekhar Borkar, William Carlson, Laura Carrington, George Chiu, Robert Colwell, William Dally, Jack Dongarra, Al Geist, Rud Haring, Jeffrey Hittinger, Adolfo Hoisie, Dean Miron Klein, Peter Kogge, Richard Lethin, Vivek Sarkar, Robert Schreiber, John Shalf, Thomas Sterling, Rick Stevens, Jon Bashor, Ron Brightwell, Paul Coteus, Erik Debenedictus, Jon Hiller, K. H. Kim, Harper Langston, Richard Miron Murphy, Clayton Webster, Stefan Wild, Gary Grider, Rob Ross, Sven Leyffer, and James Laros III. Doe advanced scientific computing advisory subcommittee (ascac) report: Top ten exascale research challenges.
- [13] Carolyn McGregor. Big data in neonatal intensive care. *Computer*, 46(6):54–59, 2013.
- [14] John Shalf, Sudip Dosanjh, and John Morrison. Exascale computing technology challenges. In *Proceedings of the 9th International Conference on High Performance Computing for Computational Science, VECPAR'10*, pages 1–25, Berlin, Heidelberg, 2011. Springer-Verlag.

ESTRATEGIAS Y ANÁLISIS ORIENTADOS AL MANEJO DE DATOS MASIVOS USANDO COMPUTACIÓN DE ALTO DESEMPEÑO

Mercedes Barrionuevo, Mariela Lopresti, Maximiliano Lucero, Natalia Miranda, Cristian Pérez
Monte, M. Antonia Murazzo(*), Fabiana Piccoli y Marcela Printista.

LIDIC- Univ. Nacional de San Luís

(*) Univ. Nacional de San Juan

San Luís, Argentina

{mbarrio, omlopres, mlucero, ncmiran, mpiccoli, mprinti}@unsl.edu.ar

RESUMEN

Los datos en nuestro universo digital han crecido de tera bytes a zetta bytes. Pero no todos los datos existentes son significativos. Un gran desafío para muchos investigadores es el descubrimiento de conocimiento a partir de un conjunto de datos muy grande en un tiempo razonable. Para lograrlo hoy se piensa en arquitecturas de naturaleza heterogéneas formadas por procesadores many y multicores.

En este trabajo se expone distintas líneas de trabajo a seguir teniendo como objetivo desarrollar técnicas de Computación de Alto Desempeño para resolver este tipo de problemas.

Palabras clave: Big Data. Computación de Alto Desempeño. GPUs. Tráfico en Redes de Computadoras.

CONTEXTO

Esta propuesta de trabajo se lleva a cabo dentro del proyecto de investigación “*Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos*” y del proyecto binacional CAPG-BA 66/13 entre la

Universidad Nacional de San Luis y la Universidad de Pernambuco, Recife, Brasil.

El proyecto de investigación se desarrolla en el marco del Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC), de la Facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis y el Centro de Informática de la UFPE.

1. INTRODUCCIÓN

El crecimiento de datos circulantes en Internet, las nuevas tecnologías y el aumento en la velocidad de transmisión de datos han dado origen al concepto de Big Data o Datos Masivos [MCJ13]. Este concepto, actualmente se define haciendo referencia a siete características: Variedad, Volumen, Velocidad, Veracidad, Viabilidad, Visualización y Valor [MCFEE12].

El conjunto de datos a manipular es tan grande y complejo que los medios tradicionales de procesamiento son ineficaces. Por lo cual es un desafío analizar, capturar, recolectar, buscar, compartir, almacenar, transferir, visualizar, etc., cantidades masivas de información, obtener conocimiento y

realizar toda su gestión en un tiempo razonable [N13].

Existen diferentes áreas o aplicaciones donde se trabaja con gran cantidad de datos, entre ellas podemos destacar la seguridad en redes de datos, recuperación de la información, evolución y parentesco de las especies, entre otras.

De acuerdo a todo lo expuesto, el procesamiento de grandes volúmenes de datos hacen que los sistemas de cómputo convencionales sean muchas veces inapropiados para lograr un procesamiento adecuado, por lo tanto una alternativa es considerar técnicas de Computación de Alto Desempeño (HPC) [YOU14], las cuales permiten realizar operaciones de cómputo intensivo y mejorar la velocidad de procesamiento; involucrando diferentes tecnologías tal como los sistemas distribuidos y los sistemas paralelos (cluster de computadoras, cloud computing, tarjetas gráficas y computadoras masivamente paralelas) [HAGER10].

Además de las técnicas de HPC, es importante la arquitectura subyacente. Hoy en día, la evolución de los sistemas de computación con multiprocesadores ha seguido dos líneas de desarrollo: las arquitecturas multicore (multi-núcleos) y las arquitecturas manycores (muchos-núcleos) [HWU08], ambos ofrecen un acceso rápido a una única memoria compartida, evitando la transferencia de datos entre distintas máquinas a través de una red. Sin embargo la cantidad de memoria disponible es limitada, una alternativa son las arquitecturas con memoria distribuida, las cuales permiten incrementar el espacio de almacenamiento (principal y secundario) aunque deben pagar el precio de la latencia de la red para llevar a cabo las comunicaciones [KAUR14]. Otra

alternativa son los sistemas híbridos, los cuales permiten combinar las características de sistemas con memoria compartida y memoria distribuida, multicores con GPU, varios sistemas en la red, entre otros. De esta manera es posible incrementar la capacidad y poder de cómputo de los sistemas computacionales permitiendo la ejecución en paralelo de múltiples procesos y threads con distintas administraciones de memoria.

Todo lo expuesto anteriormente constituye la base de nuestra motivación para investigar, verificar y poner en marcha nuevas técnicas y arquitecturas que ayuden a mejorar el procesamiento y sus tiempos de respuesta. Las técnicas de HPC serán nuestras herramientas para resolver con eficiencia cada uno de los objetivos: aplicar técnicas de HPC adecuadas para resolver problemas de datos masivos en ambientes paralelos híbridos en arquitecturas many y multicore. Como caso de estudio se presenta la búsqueda de soluciones a los diferentes los problemas planteados en la siguiente sección.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Diferentes áreas de la ciencia y la sociedad deben considerarse como casos de problemas Big Data. Estas áreas constituyen sendas líneas de investigación descritas a continuación:

- *Análisis de tráfico en términos de la seguridad de la información:*

En el campo de la detección de anomalías en redes de datos, el problema consiste en la identificación de patrones no acordes al comportamiento normal del tráfico en la red [BLMP16]. Detectar un posible ataque requiere contar con tecnologías para la clasificación del tráfico, asociando flujos de datos con las

aplicaciones que los generan. El conjunto de datos con los cuales se trabaja crece a gran velocidad, mucho mayor que su capacidad de procesamiento.

Esta línea tiene como objetivo la creación de modelos para la búsqueda de anomalías, y más precisamente, de ataques a redes de datos. Esta búsqueda consiste en la identificación de patrones que se desvían del comportamiento normal de tráfico. Tener la capacidad de tratar grandes volúmenes de información no solamente nos permite saber qué es lo que está pasando en este instante, sino también trazar patrones a lo largo del tiempo. Muchas veces es fácil pasar por alto algunos indicadores cuando analizamos información en tiempo real, sin embargo, si analizamos esa información en otros contextos y a lo largo del tiempo, quizá podamos encontrar otros significados. Lo que se pretende realizar es el procesamiento del tráfico de red, aplicarle inteligencia para obtener resultados concretos y en un tiempo cercano, de forma que la conclusión arribada no sea irrelevante, constituya la detección de un ataque y se puedan tomar decisiones rápidas.

- *Recuperación de la Información usando índices de búsquedas:*

En el caso de la recuperación de la información, el objetivo principal es satisfacer la necesidad de respuestas planteadas por un usuario en lenguaje natural, especificada a través de un conjunto de palabras claves. La información multimedia debe ser

recuperada aplicando búsquedas por similitud, por lo cual se necesitan métodos de acceso eficientes que permitan recuperar rápidamente los elementos que satisfacen los criterios de consulta. Un sistema de recuperación de información encuentra datos importantes con coincidencia parcial al patrón dado. Esta similitud, se modela utilizando una función de distancia, la cual satisface entre otras propiedades la desigualdad triangular.

Para realizar consultas en bases de datos no estructuradas o métricas se han creado modelos y algoritmos de búsqueda más generales que los correspondientes a bases de datos tradicionales [CNBY01]. El tipo de consulta en estas bases de datos se denominan consultas por similitud o proximidad, aquí los elementos son buscados considerando la cercanía de los mismos al elemento consultado. El objetivo de esta línea es utilizar índices de búsquedas aproximadas como los *permutantes* [LMPR13] para resolver consultas.

- *Construcción de Árboles Filogenéticos:*
En el área de biología evolutiva [B09, G14, Z14], el estudio de las relaciones evolutivas entre especies a partir de la distribución de los caracteres primitivos, utilizando información ADN y de morfología, es una tarea que se encuadra dentro del área de datos masivos. La idea es elaborar un árbol filogenético en el cual se muestran las relaciones evolutivas de los seres vivos entre sí, es decir, reconocer los grados de cercanía de

ancestros comunes. Este trabajo requiere contar con datos moleculares (fundamentalmente ADN) y/o morfológicos, los cuales identificarán la/s especie/s a incorporar en el árbol. Este proceso demanda gran cantidad de recursos computacionales.

El análisis filogenético requiere cálculos más complejos a medida que la cantidad de especies a catalogar crece, la cantidad de árboles a examinar en una búsqueda aumenta exponencialmente. Por esta razón es necesario contar con métodos, herramientas y técnicas de computación de altas prestaciones para datos masivos con el objeto de obtener información dentro de tiempos razonables.

Todas estas líneas de investigación tienen en cuenta la portabilidad de los desarrollos, esto se debe a que las soluciones a plantear tienen como objetivo trabajar en arquitecturas heterogéneas, aplicando técnicas de paralelismo híbrido y logrando tiempo de respuestas menores a cada uno de los problemas planteados.

3. RESULTADOS OBTENIDOS/ESPERADOS

Como objetivos de las líneas de investigación nos planteamos facilitar el desarrollo de soluciones paralelas portables, de costo predecible, capaces de explotar las ventajas de modernos ambientes de HPC a través de herramientas y “frameworks de computación” de alto nivel. Para ello será necesario proponer nuevas metodologías a ser aplicadas

en cada una de las fases del tratamiento de datos masivos.

4. FORMACIÓN DE RECURSOS HUMANOS

Los resultados esperados respecto a la formación de recursos humanos son hasta el momento el desarrollo de 4 tesis doctorales (2 concluidas), 2 tesis de maestría y varias tesinas de grado.

5. BIBLIOGRAFIA

- [B09] A. Benítez Burraco, “*Genes y lenguaje: aspectos ontogenéticos, filogenéticos y cognitivos*”. ISBN 8429110046, 9788429110043. Reverte, 2009.
- [BLMP16] Mercedes Barrionuevo, Mariela Lopresti, Natalia Miranda, Fabiana Piccoli. “*Un enfoque para la detección de anomalías en el tráfico de red usando imágenes y técnicas de Computación de Alto Desempeño*”. XXII Congreso Argentino De Ciencias de la Computación. CACIC 2016. Pp. 1166-1175. Octubre 2016, San Luis, Argentina.
- [CNBY01] Edgar Chávez, Gonzalo Navarro, Ricardo Baeza-Yates, and José Luis Marroquín. 2001. Searching in metric spaces. *ACM Comput. Surv.* 33, 3 (September 2001), 273-321.
- [G14] L. Garamszegi. “*Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice*”. ISBN 3662435500, 9783662435502. Springer, 2014.
- [HAGER10] Hager, Georg and Wellein, Gerhard. “*Introduction to high performance computing for scientists and engineers*”. Chapman & Hall/CRC computational science

- series. 1st edition, 2010. ISBN 978-1-4398-1192-4.
- [HWU08] W. Hwu, K. Keutzer, and T. G. Mattson, “*The Concurrency Challenge*” IEEE Des. Test Comput., vol. 25, no. 4, pp. 312–320, Jul. 2008.
- [KAUR14] K. Kaur and A. K. Rai, “A Comparative Analysis: Grid, Cluster and Cloud Computing,” Int. J. Adv. Res. Comput. Commun. Eng., vol. 3, no. 3, pp. 2278–1021, 2014
- [LMPR13] Lopresti, M., Miranda, N., Piccoli, F., Reyes, N. - Solving Multiple Queries through the Permutation Index in GPU. 4th International supercomputing Conference in Mexico. Colima-México. 5-8 Marzo 2013.
- [MCFEE12] A. McAfee, E. Brynjolfsson, and H. Org, “*Big Data: The Management Revolution SPOTLIGHT ON BIG DATA*” 2012.
- [MCJ13] V. Mayer-Schönberger, K. Cukier. A.I. Jurado. “*Big data: La revolución de los datos masivos*”. Turner. 2013.
- [N13] J. Needham. “*Disruptive Possibilities: How Big Data Changes Everything*”. Kindle Edition. O'Reilly Media Inc. 2013.
- [YOU14] Y. You, S. L. Song, H. Fu, A. Marquez, M. M. Dehnavi, K. Barker, K. W. Cameron, A. P. Randles, and G. Yang, “*MIC-SVM: Designing a Highly Efficient Support Vector Machine for Advanced Modern Multi-core and Many-Core Architectures*,” in 2014 IEEE 28th International Parallel and Distributed Processing Symposium, 2014, pp. 809–818.
- [Z14] Z. Yang, “*Molecular Evolution: A Statistical Approach*”. OUP Oxford, 2014.

Procesamiento de Señales y Sistemas de Tiempo Real

Modelos y Métodos Computacionales Aplicados I

Javier Giacomantone¹, Oscar Bria¹, Luciano Lorenti¹, Armando De Giusti^{1,2}

¹**Instituto de Investigación en Informática LIDI (III-LIDI)**
Facultad de Informática – UNLP
Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)
La Plata, Buenos Aires

²**CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas**

{jog, onb, llorenti, degiusti}@lidi.info.unlp.edu.ar

Resumen

Este trabajo describe una línea de investigación y desarrollo (I/D) y los resultados esperados de la misma. El objetivo principal es analizar, generar y evaluar modelos matemáticos y métodos numéricos, para abordar problemas en ingeniería y ciencias. Se estudian problemas, que por su nivel de complejidad, requieren abordajes y trabajo multidisciplinario de investigación. Los modelos abordados son dependientes del tipo de sistema estudiado, del fenómeno analizado y del área particular de ingeniería que originó el requerimiento. Determinar el tipo de sistema, el método para evaluar su rendimiento y las soluciones numéricas óptimas o sub-óptimas forma parte de los objetivos en esta línea de I/D.

Palabras Clave: métodos computacionales, análisis de sistemas, cálculo numérico, reconocimiento estadístico de patrones, análisis de imágenes, aprendizaje estadístico, desempeño y confiabilidad de sistemas.

Contexto

Esta línea de I/D forma parte del proyecto “Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real”. En particular del sub-proyecto “Modelos y

métodos computacionales. Procesamiento de señales y reconocimiento de patrones”.

1. Introducción

Modelar un problema requiere estudiar los detalles del proceso, del fenómeno, o del sistema que se pretende analizar o diseñar. Es necesaria la construcción de modelos matemáticos, métodos computacionales y técnicas numéricas que ofrezcan soluciones viables [1][2]. Las magnitudes que intervienen, escalares o vectoriales, eventualmente, cambian en el tiempo o espacio. Los sistemas estudiados pueden ser lineales o no lineales requiriendo modelos sofisticados [3][4]. Determinar si una solución es viable con fundamento científico, es una tarea compleja y dependiente del problema particular analizado [5]. Este proyecto tiene como primer objetivo analizar y proponer modelos computacionales, métodos y soluciones particulares derivadas de los mismos. Alcanzar los objetivos anteriores requiere estudiar los fundamentos que subyacen a cada modelo, evitando soluciones, que por su nivel de encapsulamiento, limiten una verdadera comprensión y abordaje científico de las mismas [6]. El tipo de problema de interés en esta línea de I/D requiere la integración de soluciones de tres áreas, ciencias de la computación, matemáticas aplicadas y un área de ingeniería o ciencia básica en particular. La evaluación

de rendimiento es un aspecto fundamental para poder validar las soluciones propuestas o los modelos analizados [7]. Por lo tanto, otro aspecto fundamental de esta línea de I/D es el estudio de las métricas y paradigmas de desempeño en sistemas específicos.

En la sección 2 se presenta un breve resumen de los temas de I/D específicos en el período actual. La sección 3 enumera resultados obtenidos y esperados. Finalmente, la sección 4 resume los objetivos con respecto a la formación de recursos humanos en el contexto de esta línea de I/D.

2. Líneas de Investigación

2.1 Reconocimiento de patrones

El trabajo actual lo podemos clasificar en tres tópicos principales bien diferenciados correspondientes a sub-disciplinas dentro de reconocimiento estadístico de patrones [8][9]. La primera sub-disciplina es clasificación supervisada donde el énfasis de nuestro trabajo se centra en el estudio de métodos de clasificación basados en núcleos dispersos, en particular máquinas de soporte vectorial [10][11].

La segunda es clasificación no supervisada donde la principal línea de trabajo son las técnicas de agrupamiento. Actualmente con énfasis en detección de valores atípicos y métodos basados en teoría espectral de grafos [12][13]. El tercer tópico de fundamental importancia es el de reducción de dimensión en particular selección de características [14].

2.2 Análisis de Imágenes

Se estudian métodos de segmentación estáticos y modelos deformables. En particular en imágenes 2 ½ D obtenidas por cámaras de tiempo de vuelo, resonancia magnética funcional y otras modalidades con estructuras de datos similares [15]. El objetivo principal es mejorar la calidad de los descriptores obtenidos en función de su impacto en el sistema de clasificación [16][17]. En el caso de

modelos deformables es posible estudiar el comportamiento temporal y medir magnitudes indirectamente. Se analizan métodos de generación de características a partir de señales en general y de imágenes digitales en particular de rango e intensidad. Se abordan sistemas de análisis de pseudo-imágenes a partir de campos vectoriales, series temporales, una modalidad particular de imagen o fusión de modalidades.

2.3 Desempeño de Sistemas de Posicionamiento, Navegación y Localización.

En los sistemas de posicionamiento, de navegación y de localización [18][19], el concepto de desempeño excede al habitual que está limitado a la calidad nominal de la estimación de ubicación y eventualmente a la confiabilidad [20][21]. En estos sistemas deben considerarse además los parámetros de integridad y continuidad que le garanticen al usuario que la información proporcionada por el sistema es correcta para que una operación crítica pueda realizarse en forma segura [22][23].

La integridad y la continuidad dependen en gran medida de la aplicación y del entorno específico y su aseguramiento afecta a otro parámetro de desempeño del sistema que es la disponibilidad [24].

Esta línea de trabajo se avoca al estudio de problemas puntuales de desempeño en los sistemas mencionados, utilizando criterios y métodos diversos de modelado, procesamiento y análisis [25][26].

3. Resultados y Objetivos

3.1 Resultados publicados recientemente

- Se estudiaron y propusieron métodos para detección en series temporales de fMRI [27][28].
- Se desarrollaron métodos de segmentación de imágenes de rango y supresión del plano de fondo [29][30][31][32].

- Se analizaron y propusieron alternativas para el agrupamiento de objetos de interés en video [33].
- Se estudió el desempeño de un método de exclusión de satélites en un sistema de ayuda a la aeronavegación basado en GNSS [34].
- Se propuso y se presentaron resultados experimentales de un método de aprendizaje basado en problemas para aritmética computacional [35].

3.2 Objetivos generales

- Desarrollar modelos y optimizar algoritmos particulares de clasificación supervisada y no supervisada.
- Evaluar métodos de análisis de desempeño y su aplicación sobre los clasificadores y conjuntos de datos particulares.
- Evaluar la monitorización de la integridad de los sistemas de ayuda a la navegación aérea basados en sistemas GNSS (Global Navigation Satellite Systems).
- Estudiar métodos de selección y extracción de características.
- Investigar modelos y métodos computacionales en procesamiento y análisis de imágenes.
- Promover la interacción con otros grupos y líneas de I/D resultando en un mecanismo de permanente consulta y transferencia.

4. Formación de Recursos Humanos

Se dictan asignaturas optativas en Ingeniería en Computación y cursos de postgrado que tienen por objetivo formar alumnos en temas específicos y fundamentos. Los alumnos tienen la posibilidad de realizar trabajos de investigación asociados a esta línea de trabajo.

BIBLIOGRAFIA

1. Zalizniak V. Essentials of Scientific Computing – Numerical Methods for Science and Engineering. Woodhead Publishing, 2008.
2. Juergen G. Coupled Systems: Theory, Models, and Applications in Engineering. CRC, 2014.
3. Torokhti A., Howlett P. Computational Methods for Modelling of Nonlinear Systems. Elsevier, 2007.
4. Canuto C., et al. Spectral Methods. Evolution of Complex Geometries and Applications to Fluids Dynamics Scientific Computation. Springer, 2007.
5. Ciurpina G. Scientific Computing in Electrical Engineering Springer, 2007.
6. Gustafsson B. Fundamentals of Scientific Computing Springer, 2011.
7. Aslak T., et al. Elements of Scientific Computing. Springer, 2010.
8. Fukunaga K. Introduction to Statistical Pattern Recognition. Second Edition. Academic Press, 1990.
9. Devijer P, Kittler, J. Pattern Recognition: theory and applications. Springer, 1986.
10. Corte C, Vapnik V, Support vector networks. Machine Learning v.20, pp.273-297, 1995.
11. Vapnik, V. The Nature of Statistical Learning Theory. N. Y. Springer, 1995.
12. Von Luxburg U. A Tutorial on Spectral Clustering. Statistics and Computing, 17(4), 2007.
13. Shi J., Malik J. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(8), 888-905, 2000.
14. Aytug H. Feature selection for support vector machines using Generalized Benders Decomposition. European Journal of Operational Research, 244(1), 210-218, 2015.
15. Kim H.Y., Giacomantone J. O., Cho, Z. H. Robust Anisotropic Diffusion to Produce Enhanced Statistical Parametric Map, Computer Vision and Image Understanding, v.99, pp.435-452, 2005.
16. Han Y., Feng X., Baciú G. Variational and PCA based natural image segmentation. Pattern Recognition 46, pp. 1971-1984, 2013.

17. Li S., Fevens L., Krzyzak A., Li S. Automatic Clinical Image Segmentation Using Pathological Modelling, PCA and SVM, MLDN, LNAI 3587 pp.314-324, 2005.
18. Partap Misra, Per Enge. Global Positioning System: Signals, Measurements and Performance, Ganga-Jamuna Press, 2010.
19. Hakan Koyuncu, Shuang Hua Yang. A Survey of Indoor Positioning and Object Locating Systems Indoor Positioning System, International Journal of Computer Science, 2010.
20. Petevelo Mark. Quantifying the performance of Navigation Systems and Standars for assisted-GNSS, Inside GNSS, 2008.
21. Morurikis A., Roumeliotis S. Performance Analysis of Multirobot Cooperative Localization, IEEE, 2005.
22. Murphy T., et. al., Fault Modeling for GBAS Airworthiness Assessments, Navigation, 2012.
23. Cosmen-Schortmann J., Azaola-Sáenz, Martínez-Olagüe M. A., Toledo-López M., Integrity in Urban and Road Environments and its use in Liability Critical Applications, IEEE, 2008.
24. Shuo-Ju Yeh1, Shau-Shiun, GBAS airport availability simulation tool, GPS Solutions, 2015.
25. Pengfei Duan, Maarten Uijy De Haa. Flight Test Results of a Measurement-Based ADS-B System for Separation Assurance, Navigation, 2013
26. Sam Pullen, Todd Walter, Per Enge. SBAS and GBAS Integrity for Non-Aviation Users: Moving Away from "Specific Risk," International Technical Meeting of The Institute of Navigation, 2011.
27. Giacomantone J., Tarutina T. Diffuse Outlier Detection Technique for Functional Magnetic Resonance Imaging. Computer Science and Technology Series. XVI Argentine Congress of Computer Science Selected Papers. pp. 255-265, 2011.
28. Giacomantone J., De Giusti A. Detección de áreas de interés bajo la hipótesis de relación espacial de voxels activados en fMRI. XX Congreso Argentino de Ciencias de la Computación. San Justo. Argentina, 2014.
29. Lorenti L., Giacomantone J. Segmentación espectral de imágenes utilizando cámaras de tiempo de vuelo. XI Workshop Computación Gráfica, Imágenes y Visualización. pp. 430-439. Mar del Plata, Argentina, 2013.
30. Lorenti L., Giacomantone J. Time of flight image segmentation through co-regularized spectral clustering. Computer Science & Technology Series. XX Argentine Congress of Computer Science. Selected papers. La Plata, Editorial de la Universidad Nacional de La Plata, 2015.
31. Giacomantone J., et al. Supresión del plano de fondo en imágenes de tiempo de vuelo. VII Workshop Procesamiento de Señales y Sistemas de Tiempo Real, 2016.
32. Lorenti, L., Giacomantone, J., Bria, O. N., De Giusti, A. E. (2017). Fusión de información de geometría e intensidad para segmentación de imágenes TOF. XXIII CACIC. La Plata, Argentina, 2017.
33. Lorenti L., Giacomantone J., De Giusti A. Agrupamiento de trayectorias vía clustering espectral incremental. XXII CACIC, pp. 222-231, 2016.
34. Bria, O., Giacomantone, J., Lorenti, L., Excluding Ionosphericly Unsafe Satellite Geometries in GBAS CAT-I. XXII CACIC. La Plata, Argentina, 2017.
35. Giacomantone, J., Bria, O., Proactive Independent Learning Approach: A case study in computer arithmetic. XXII CACIC. La Plata, Argentina, 2017.

ENERGÍA EFICIENTE

Bruno Boyle, Estefano Biasin & Pedro Lopez.

*Facultad de Tecnología Informática, Universidad Abierta Interamericana. Av. Ovidio Lagos 944,
Rosario, Santa Fe, Argentina.*

brunoboyle@gmail.com

estefanobiasin@gmail.com

RESUMEN

El objetivo del proyecto es desarrollar un dispositivo (prototipo), que permita a cualquier persona tener noción del consumo eléctrico domiciliario. Mediante tecnología Arduino y sensores ubicados en distintas partes del circuito eléctrico, compatible con Arduino, capaz de cuantificar magnitudes en tiempo real, como ser consumo, valores de tensión y corriente. A partir de dicha acción y con la conexión del dispositivo vía internet, mediante un servidor, se llevará a cabo el procesamiento de los datos, brindando información útil para el usuario, como puede ser: consumos y niveles de tensión, notificaciones y alertas configurables por el usuario y una estimación monetaria del consumo. Esto será visualizable desde cualquier dispositivo con conexión a internet. Haciendo así un uso consciente y responsable de la energía eléctrica.

Palabras clave - IoT, web, arduino, consumo, información.

CONTEXTO

La necesidad de poder cuantificar el consumo eléctrico, la responsabilidad de utilizar dicha energía eficientemente, así como saber cuándo los niveles de tensión no son los pre establecidos o cuando la instalación eléctrica está sobrecargada, lo cual genera, que los electrodomésticos, no funcionen de una forma

correcta o eficiente, obligan a la comunidad científica y técnica a realizar tareas de investigación y pruebas para poder responder a dicha necesidad con una solución viable, económica y eficaz

A continuación, se hace referencia a los trabajos que funcionaron como punto de partida de este proyecto

El trabajo número 1, podemos encontrar una solución no tecnológica. Es un proceso meramente manual, el cual consta de estimar el consumo de cada electrodoméstico conectado a la instalación eléctrica del domicilio y aproximar su tiempo y simultaneidad de uso para poder tener una noción de la carga y consumo en el circuito, en tiempo real. A partir de dicha estimación, mediante otro proceso manual como lo es realizar el cálculo de consumo por unidad de tiempo y cuantificación económica, obtener el importe de lo consumido. Esta solución no satisface de forma completa la necesidad anteriormente expuesta ya que, entre otros motivos, no cumple con las pautas establecidas como lo son la viabilidad y eficacia de la misma.

El trabajo número 2, encontramos un dispositivo electrónico, el cual se conecta a un tomacorriente domiciliario y ofrece un tomacorriente de salida. El mismo posee un indicador que muestra el consumo del electrodoméstico conectado a dicho tomacorriente. El mismo es de un costo

económico bajo, pero carece de prestaciones y personalización.

Tomando como referencia los trabajos anteriores expuestos, a continuación, se presenta una tabla comparativa buscando los ítems más significativos y referenciales para llevar a cabo el proyecto.

Tabla 1. Comparación de dispositivos y trabajos anteriores con el proyecto actual.

Trabajos	C1	C2	C3	C4
Estimación a través de técnicas manuales de NO tecnología		X		X
Dispositivo electrónico convencional	X		X	X
Trabajo actual	X	X	X	X

C1: Alta tecnología. C2: Personalización. C3: Automatización. C4: Bajo costo.

INTRODUCCIÓN

Desde el descubrimiento de la energía eléctrica existió el debate de cómo hacer uso de la misma. Si nos remontamos a los años 1880, donde tiene lugar la llamada “Guerra de las corrientes” Nikola Tesla y Thomas Edison, se debatían sobre la producción y distribución de corriente eléctrica.Cuál de las corrientes eléctricas, si la alterna o la continua, sería más costosa, cuál más fácil de producir y distribuir y cual más segura para las personas, entre otros puntos. Hoy en día, con todas estas incógnitas ya resueltas en su gran mayoría, nos enfrentamos a otro debate, que es la utilización de la energía eléctrica en forma eficiente y las energías renovables. En Argentina para el año 2013 solo el 1% de

la energía eléctrica generada era renovable. El gobierno de la nación Argentina, promulgo la ley de “Régimen de Fomento a la Generación distribuida de Energía Renovable integrada a la Red Eléctrica Pública”, la cual entre otras ítems enuncia, que para el 2017 el 8 % de la energía debería ser renovable, llegando a un 20% para el año 2020. Por lo cual se puede evidenciar que el paradigma respecto a las energías ha cambiado, ya no se considera a las mismas como una fuente inagotable. Se tienden a generar energías renovables y a hacer un uso eficaz de las mismas.

Ser eficientes a la hora de consumir energía quiere decir utilizar menos energía por un mismo servicio.

Con la revolución propuesta por lo teléfonos inteligentes, internet a llegado a cualquier aspecto de nuestras vidas, pudiendo, entre otras cosas, cuantificar variables y magnitudes, que antes no creíamos así. Siguiendo esta corriente, hoy en día se encuentra en auge la “internet de las cosas”, también conocido como IoT (“internet off things”).El concepto establece la interconexión digital de objetos a internet. Un ámbito en donde todavía dichas tendencias no incursionaron es en la energía eléctrica domiciliaria.

Situación Problemática

La energía eléctrica se volvió un recurso básico para cualquier persona, el prescindir de la misma trae aparejadas problemáticas que podrían ser de índole vital y social

Lo inconvenientes generados a partir del derroche de energía eléctrica o el exceso de consumo, pueden producir cortes en el suministro de dicho servicio, especialmente en época de temperaturas elevadas, donde la demanda es muy alta. En estas épocas también podemos afirmar que los niveles de

tensión no son los ideales, por lo cual, para que los electrodomésticos puedan funcionar y mantener su potencia, se incrementa la corriente, pudiéndose producir, a causa de eso, una sobre carga en el circuito eléctrico hogareño, siendo totalmente desconocido por el usuario, hasta el momento que entra en acción alguna de las protecciones del circuito, pero dejando momentáneamente el mismo sin energía.

Objetivos específicos

- Analizar antecedentes sobre la aplicación de este tipo de dispositivos en hogares
- Diseñar un dispositivo (prototipo) con Sistema Arduino que permite visualizar en tiempo real, el consumo eléctrico, parámetro de tensión y corriente, cuantificar dicho consumo en tarifa.
- Proporcionar al usuario una herramienta, para tener control sobre su consumo eléctrico y estado del circuito, evitando posibles cortes de energía y gastos económicos.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proceso de monitoreo en forma continua del circuito eléctrico del hogar es una innovación en lo que respecta al uso eficiente de la energía. Permitiendo obtener información en forma rápida, concisa y real de los parámetro anteriormente mencionados, Para que este proyecto brinde una solución, viable, económica y eficaz, deberá cumplir con las siguientes características:

- Fácil uso
- Fácil Instalación
- Entorno Gráfico Amigable
- Opciones de personalización

- Escalabilidad
- Seguro

Lo que se propone como solución es una integración de distintas tecnologías para poder llevar a cabo un dispositivo con sensores que brinde al usuario la posibilidad de saber su consumo eléctrico en tiempo real y expresarlo en forma monetaria, así como también notificar de excesos de consumos, de sobrecargas del circuito o de baja tensión. Dicho dispositivo permitirá, a través de una aplicación WEB una interfaz accesible desde cualquier dispositivo con acceso a internet. Pudiendo configurar parámetro y alertas y visualizar esta información de forma fluida.

Este dispositivo se compone de las siguientes herramientas:

Módulo Central:

- **Arduino UNO**
- **Ethernet Shield**

Sensores:

- **Optoacoplador:** Encargado de medir la tensión.
- **Sensor de efecto hall:** Encarga de medir la corriente.

Servidor:

- **Servidor linux:** Encargado de recibir y procesar los datos brindados por el módulo central y luego disponibilizar a cualquier cliente web.

Dispositivo Cliente:

- **Cualquier dispositivo con acceso a internet:** Encargado de visualizar la información del consumo eléctrico en tiempo real.



Figura 1: Arduino Uno R3.

El Arduino Uno R3 es una placa electrónica de las muchas que tiene Arduino y con la que es muy fácil introducirse en el mundo de la programación electrónica, Arduino es una plataforma de código abierto (open-source) lo que permite realizar proyectos y modificaciones tanto de hardware como de software a cualquier persona sin ningún problema.

Características técnicas:

- Microcontrolador ATmega328.
- Voltaje de entrada 7-12V.
- 14 pines digitales de I/O (6 salidas PWM).
- 6 entradas análogas.
- 32k de memoria Flash.
- Reloj de 16MHz de velocidad.



Figura 2: Arduino ethernet shield.

El Arduino ethernet shield nos da la capacidad de conectar un Arduino a una red ethernet. Es la parte física que implementa la pila de protocolos TCP/IP. Está basada en el chip ethernet Wiznet W5100. El Wiznet

W5100 provee de una pila de red IP capaz de soportar TCP y UDP. Soporta hasta cuatro conexiones de sockets simultáneas. Usa la librería Ethernet para leer y escribir los flujos de datos que pasan por el puerto ethernet.

El shield provee un conector ethernet estándar RJ45. La ethernet shield dispone de unos conectores que permiten conectar a su vez otras placas encima y apilarlas sobre la placa Arduino.

Arduino usa los pines digitales 10, 11, 12, y 13 (SPI) para comunicarse con el W5100 en la ethernet shield. Estos pines no pueden ser usados para e/s genéricas.

El botón de reset en la shield resetea ambos, el W5100 y la placa Arduino.

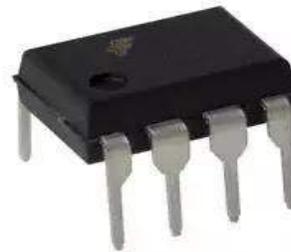


Figura 3: Optoacoplador HCPL-3700

El optoacoplador HCPL-3700 con detección umbral de tensión/corriente consiste en un LED AlGaAs conectado a una entrada de detección de umbral buferizada el cual es ópticamente acoplado a una salida Darlington alta ganancia. El chip de memoria intermedia de entrada es capaz de controlar los niveles de umbral en un amplio rango de voltajes de entrada con una sola resistencia. La salida es compatible con TTL y CMOS.

Características:

- Entrada AC/DC
- Sensibilidad de voltaje programable
- Compatibilidad a nivel lógico
- Umbral de sobre temperatura garantizado (0°C a 70°C)

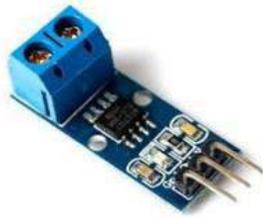


Figura 4: sensor ACS712

El sensor ACS712 de Allegro, es un sensor de corriente por efecto hall, que provee una solución económica y precisa para medir corriente en AC o DC, ya sea en ambientes industriales o comerciales. Este Sensor funciona transformando un campo magnético surgido de el paso de la corriente por un alambre de cobre interno en el sensor, y convirtiendo este campo en un voltaje variable. Esto significa que a mayor cantidad de corriente que tengamos, mayor voltaje vamos a tener en un pin. Este sensor viene en 3 modelos distintos: ACS712ELCTR-05B-T que mide hasta 5A, el ACS712ELCTR-20A-T que mide hasta 20A y el ACS712ELCTR-30A-T que mide hasta 30A. Las diferencias entre cada uno de los modelos es que las variaciones de voltaje en su pin de salida es siempre la misma, por ende, para cualquier modelo su salida analógica variara entre 0 y 5v

RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos en la puesta en marcha de dicho dispositivo presentan grandes beneficios, ya que podemos ver en tiempo real el costo económico de todos los electrodomésticos funcionando y saber si estamos dentro de los parámetros para los cuales el circuito fue diseñado.

Podemos concluir que este dispositivo junto con la modificación en las rutinas del usuario,

el reemplazo de alumbrado y equipos eléctricos viejos, por alternativas más eficiente, pueden proporcionar a medio y largo plazo una disminución en el consumo eléctrico lo cual implica un uso consciente y eficiente.

FORMACIÓN DE RECURSOS HUMANOS

El proyecto está dirigido por los Ingenieros Pedro López y Pablo Audoglio, ambos docentes de Universidad Abierta Interamericana. Además, integran el proyecto los alumnos Bruno Boyle y Estefano Biasin, ambos cursando el último año de Ingeniería en Sistemas de Información en la UAI Sede Lagos, Rosario, Santa Fé, Argentina.

BIBLIOGRAFÍA

- [1] ¿Qué significan las especificaciones técnicas del Arduino?
- Fuente panamahitek-
Disponible en:
<http://panamahitek.com/especificaciones-tecnicas-del-arduino/>
- [2] Getting Started with the Arduino Ethernet Shield
- Fuente [arduino](http://arduino.cc)
Disponible en:
<https://www.arduino.cc/en/Guide/ArduinoEthernetShield>
- [3] Arduino UNO r3
Fuente [arduino](http://arduino.cc)
Disponible en:
<https://www.arduino.cc/en/Guide/Arduino>
- [4] Tecnología Arduino – Fuente Arduino
Disponible en: <http://arduino.cl/que-es-arduino/>
- [5] HCPL-3700 Optoacoplador AC/DC Interfaz Lógica – Fuente Microjpm
Disponible en:
<https://www.microjpm.com/products/hcpl-3700-optoacoplador-ac-dc-a-interfaz-logica/>
- [6] Medidor de corriente con el sensor ACS712 – Fuente Punto flotante
Disponible en:
<https://puntoflotante.net/SENSOR-DE-CORRIENTE-ACS712.htm>

Robótica y Redes de Sensores en Sistemas de Tiempo Real

Fernando Romero, Mariano Mendez, Diego Encinas, Armando De Giusti¹,
Santiago Medina, Martín Pi Puig, Horacio Villagarcía, Juan Manuel Paniego,
Fernando G. Tinetti²

Instituto de Investigación en Informática LIDI (III-LIDI)³
Facultad de Informática – Universidad Nacional de La Plata

¹ CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

² Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

³ Centro Asociado de la CIC

{fromero, mmendez, dencinas, degiusti, smedina, mpipuig, hvw, jmpaniego, fernando}@lidi.info.unlp.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio y desarrollo de Sistemas de Tiempo Real (STR), en particular los que incumben a robots del tipo móvil y redes de sensores. Se llevan a cabo experimentos con microrrobots y con sistemas de microcontroladores. A fin de corroborar hipótesis se implementa en escenarios de sistemas reales y simulados.

Contexto

Esta línea de Investigación forma parte del proyecto "Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real " del Instituto de Investigación en Informática LIDI acreditado por la UNLP.

Palabras Claves: Tiempo Real, Simulación, Sistemas Embebidos, Comunicaciones, redes de Sensores, Robots, Microcontroladores.

1. Introducción

Los Sistemas de Tiempo Real (STR) tienen como característica principal que no solo elaboran una respuesta correcta ante una entrada determinada, sino que la misma debe ocurrir dentro de un plazo [5] [6] [11] [12] [16] [17], Ello es debido a que deben respetar los tiempos del sistema en el cual funcionan. Interaccionan con un entorno físico a partir de tomar datos del mismo a través de sensores y elaborar respuestas, que muchas veces llega al mismo entorno a través de actuadores. Es frecuente que estos STR realicen tareas emulando trabajo humano, lo que se conoce como robótica. Un tipo de robot de gran desarrollo son los móviles, que presentan capacidad de desplazamiento, basados en una plataforma con un sistema locomotor. Para elaborar su recorrido se utilizan diferentes tipos de sensores, entre ellos codificadores que miden el desplazamiento, y diferentes sistemas de radio que ayudan a calcular su posición. De ellos el más conocido es el GPS (Geo Posicionamiento Satelital). La señal de los satélites no es capaz de llegar con la intensidad necesaria a los espacios interiores. Además, si llegara, habría que

incluir mapas de edificios entendibles por un dispositivo móvil. Debido a ello, GPS debe complementarse con IPS (sistema de posicionamiento en interiores, en inglés indoor positioning system). Dentro de estos sistemas IPS tenemos los llamados fingerprinting, basados en WiFi o cualquier tipo de señal que pueda ser captado por el móvil: WiFi, campo magnético, GPS. Las señales se utilizan para obtener una referencia sobre un mapa, de manera similar a un GPS. Este mapa incluye un conjunto de puntos de posición, asociados a la lista de señales, por ejemplo redes WiFi, captadas por el móvil, con la intensidad de cada una. Para el posicionamiento, se compara el valor que mide el móvil a posicionar con el mapa de referencia.

Dentro de los robots móviles encontramos dos tipos: los que elaboran la respuesta en forma puramente computacional, y los que emplean intervención humana [21].

Dos características a estudiar en estos sistemas móviles son el consumo y la latencia en la elaboración de las respuestas frente a señales externas.

Se experimenta sobre diferentes SOTR y sobre simulaciones de los mismos. Además, se disponen de diversas plataformas de hardware y software para la construcción de sistemas de diferente grado de complejidad. Estos pueden estar basados en microcontroladores [9] que soporten o no un SOTR. Entre las placas de desarrollo utilizadas se encuentran Arduino, Raspberry PI, NodeMcu [27] y CIAA [19], entre otras.

Una característica fundamental de los SOTR es tener un alto grado de fiabilidad, por lo que el estudio de la detección y control de condiciones de falla es un aspecto de gran importancia [2] [14].

En las diferentes implementaciones que se llevan a cabo en esta investigación, se trabaja con mini-robots armados por alumnos y Khepera [18], sobre diferentes SOTR (Linux RT-Preempt, FreeRTOS, MQX, OSEK-OS, etc.) [8].

Se estudian redes de sensores basados en los protocolos cableados RS485, SPI, I2C, CANBUS [20] y MODBUS [1] e inalámbricos [28] [29] [30].

Se puede disminuir la complejidad y la probabilidad de errores en el diseño de sistemas, desarrollando una simulación específica a través de la utilización de diferentes frameworks [24]. En particular, con CANBUS se han realizado algunas implementaciones reales y simulaciones con la herramienta Proteus [25].

2. Líneas de Investigación y Desarrollo

Se plantean como temas de estudio:

- Sistemas robóticos con intervención humana [4] en el lazo de control a través de acciones e interfaces no convencionales. Son componentes de estos sistemas: el modelo de reconocimiento de comandos gestuales [22] [23], las interfaces, el modelo de observación, los modelos dinámicos y de realimentación de la máquina, el modelo de planificación y, eventualmente, los modelos de realimentación para el operador humano y el modelo de actuación humana (para interpretar acciones y distracciones).
- Verificación y validación del hardware por medio de simulaciones que permitan predecir posibles comportamientos y en consecuencia generar mayor eficiencia de diseño [7] [8].
- Implementación y simulación de redes cableadas de sensores interactuando con un computador. Se utiliza MODBUS, RS485 o CANBUS como

protocolo de comunicaciones, tanto a nivel físico como de simulación [1] [15].

- Construcción de redes inalámbricas específicamente orientadas a sensores. En principio, se implementarán experimentos para caracterizar estas redes en términos de métricas como latencia y ancho de banda para el caso de rendimiento, distancia (alcance), confiabilidad (pérdida de paquetes), etc.
- Evaluación de redes específicamente diseñadas para distancias mayores a los estándares de WiFi (ej: LoRa) [26]
- Odometría a través de robots Khepera [13] [3] y otros de producción propia.
- IPS a partir de fingerprinting WiFi. Se enfoca principalmente en obtener vehículos autónomos [10] que puedan circular en un entorno de autopistas inteligentes y con capacidad de estacionamiento.

3. Resultados y Objetivos

Se han desarrollado tareas sobre los temas antes expuestos tales como:

- Experimentos en Odometría.
- Medición de consumo de diferentes microcontroladores en distintas condiciones de uso y diferentes SOTR.
- Estudio de plataformas de hardware: Arduino, Intel Galileo, CIAA, Freescale Kinetis, Raspberry Pi.
- Construcción y estudio de redes de sensores cableadas, empleando CANBUS, MODBUS y RS485.
- Construcción y estudio de redes de sensores inalámbricas.

4. Formación de Recursos Humanos

En base a estos temas se desarrollan trabajos de varios alumnos en el marco de

la Convocatoria a Proyectos de Desarrollo e Innovación de la Facultad de Informática de la UNLP. Además, se encuentran en desarrollo y concluidas varias tesinas de grado de alumnos de la Licenciaturas de Informática y Sistemas, como así también Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Computación. En cuanto a postgrado, se encuentra en desarrollo un trabajo final de especialización y una tesis de Magister.

5. Referencias

- [1] Jordi Bartolomé "El protocolo MODBUS", 2011. En <http://www.tolaemon.com/docs/modbus.htm>
- [2] Andersen, B. L. "Method of detecting systemic fault conditions in an intelligent electronic device." U.S. Patent 6,434,715, issued August 13, 2002.
- [3] Azizi, F., N. Houshang. "Mobile robot position determination using data from gyro and odometry." In Electrical and Computer Engineering, 2004. Canadian Conference on, vol. 2, pp. 719-722. IEEE, 2004.
- [4] Bekey, George A. Robotics: state of the art and future challenges. Imperial College Press, 2008.
- [5] Burns, A, A, Wellings. "Real-Time Systems and Programming Languages: Ada, Real-Time Java and C/Real-Time POSIX", Addison-Wesley Educational Publishers Inc., 2009.
- [6] Buttazzo, G. C., "Hard RealTime Computing Systems", Third edition, Springer, 2011.
- [7] Eickhoff, J., Simulating Spacecraft Systems, Springer, 2009.
- [8] "FreeRTOS - market leading RTOS (real time operating system) for embedded systems supporting 34 microcontroller architectures". <http://www.freertos.org/>.

- [9] Jenkins, T., I. Bogost. "Designing for the internet of things: prototyping material interactions." In CHI'14 Extended Abstracts on Human Factors in Computing Systems, pp. 731-740. ACM, 2014.
- [10] Jones, J. L., A. M. Flynn, Bruce A. Seiger. Mobile robots: inspiration to implementation. Vol. 2. Wellesley MA: AK peters, 1999.
- [11] Kopetz. H., "Real-Time Systems, Design Principles for Distributed Embedded Applications". Second Edition. Springer. 2011.
- [12] Liu, J. W. S. Liu, "Real Time Systems", Integre Technical Publishing Co., Inc., 2000
- [13] Rekleitis, I. M., G. Dudek, E. E. Milios. "Multi-robot exploration of an unknown environment, efficiently reducing the odometry error." In International Joint Conference on Artificial Intelligence, vol. 15, pp. 1340-1345. Lawrence Erlbaum Associates Ltd, 1997.
- [14] Rohani, A., H. R. Zarandi. "An analysis of fault effects and propagations in AVR microcontroller ATmega103 (L)." In Availability, Reliability and Security, 2009. ARES'09. International Conference on, pp. 166-172. IEEE, 2009.
- [15] Introduction to the Controller Area Network (CAN) Texas Instrument Application Report SLOA101A–August 2002–Revised July 2008.
- [16] Silberschatz, A., P. B. Galvin, G. Gagne, Operating System Concepts, 8th Edition, ISBN : 978-0-470-12872-5, Wiley, 2009.
- [17] PHILLIP A. LAPLANTE, SEPPO J. OVASKA. REAL-TIME SYSTEMS DESIGN AND ANALYSIS Tools for the Practitioner Fourth Edition. A JOHN WILEY & SONS, INC., PUBLICATION. IEEE PRESS. 2012.
- [18] <http://ftp.k-team.com/khepera/documentation/Kh2IRAN.pdf>
- [19] <http://www.proyecto-ciaa.com.ar/devwiki/doku.php?id=desarrollo:edu-ciaa:edu-ciaa-nxp>
- [20] <http://www.can-cia.de/can-knowledge/can/can-fd/>
- [21] Chi-Pang Lam and Shankar Sastry, A POMDP Framework for Human-in-the-Loop System, University of California at Berkeley.
- [22] Mitra S., Acharya T." Gesture recognition: A survey". IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol 37(3). pp 311–324. 2007.
- [23] Murthy G. R. S., Jadon R. S. "A Review of Vision Based Hand Gestures Recognition". International Journal of Information Technology and Knowledge Management. Vol.2-2. pp 405-410. 2009.
- [24] D. Black, SystemC: From the Ground Up. Second Edition, Springer, 2010.
- [25] Proteus. <https://www.labcenter.com>. 2017
- [26] LoRa <https://www.lora-alliance.org/> 2017
- [27] NodeMcu <http://www.nodemcu.com/> 2017
- [28] Akyildiz, Ian F., and Mehmet Can Vuran. "Wireless sensor networks" Vol. 4. John Wiley & Sons, 2010.
- [29] Lewis, Franck L."Wireless sensor networks." Smart environments: technologies, protocols, and applications 11 (2004): 46.
- [30] Raghavendra, Cauligi S., Krishna M. Sivalingam, and Taieb Znati, eds. "Wireless sensor networks" Springer, 2006.

Estación Terrena de Adquisición de Señales de Satélites no Geoestacionarios

Facundo Busano[†], Marco Miretti, Emmanuel DAVIS, Emanuel Bernardi,
Hugo Pipino, Sergio Felissia, Gastón Peretti
Grupo de Investigación y Desarrollo Electrónico (GIDE)
Universidad Tecnológica Nacional
Facultad Regional San Francisco (UTN-FRSFCO)
San Francisco, Córdoba, Argentina
[†]facundobusano@gmail.com

RESUMEN

El creciente desarrollo de los nanosatélites ha conducido a empresas privadas y universidades a involucrarse fuertemente en la industria aeroespacial. Pese a que cada operador de nanosatélite suele poseer su propia estación terrena de comunicación y control, la reducida ventana de tiempo disponible para establecer la comunicación, por tratarse de órbitas terrestres bajas, hace necesaria una red global de recolección de datos. Es por ello, que mediante este proyecto se pretende desarrollar un sistema de posicionamiento de múltiples antenas que reciben señales (UHF, VHF y S-Band) provenientes de satélites no geoestacionarios, entre ellos los nanosatélites o CubeSat. Para luego, emplear dichas señales, compuestas de datos de telemetría e imágenes, en el monitoreo de la salud de cultivos, la cuantificación de superficies inundadas y el análisis climático de nuestro entorno.

Por último, además de adquirir, procesar y utilizar la información recolectada, el proyecto contempla la colaboración con los operadores ya establecidos en distintas organizaciones a nivel mundial.

Palabras clave: comunicación, sistema de control, procesamiento de imágenes, nanosatélites, posicionamiento.

CONTEXTO

Esta propuesta de investigación se enmarca dentro de las áreas Control Automático de

Sistemas y Telecomunicaciones del Departamento de Ingeniería Electrónica de la Universidad Tecnológica Nacional, Facultad Regional San Francisco. Además, las actividades involucradas en este proyecto se llevarán a cabo por los integrantes del Grupo de Investigación y Desarrollo Electrónico (GIDE), el cual se compone de becarios alumnos, docentes investigadores y becarios doctorales.

1. INTRODUCCIÓN

En las últimas décadas, más precisamente a partir del año 1999, la industria aeroespacial ha abordado, en conjunto con universidades y empresas privadas, cientos de proyectos de desarrollo de nanosatélites (término que hace referencia a satélites artificiales cuya masa se encuentra entre 1 y 10 kg), también denominados CubeSat. Éstos, surgieron de la unión de esfuerzos entre la Universidad Politécnica Estatal de California (San Luis Obispo, EEUU) y el Laboratorio de Desarrollo de Sistemas Espaciales de la Universidad Stanford (Palo Alto, EEUU) con el objetivo de proporcionar un estándar para el diseño de nanosatélites que reduzcan los costos y el tiempo de desarrollo, incrementando la accesibilidad al espacio y la frecuencia de lanzamientos.

En la actualidad, el proyecto CubeSat es una colaboración internacional de más de 100 universidades, escuelas secundarias y empresas privadas que desarrollan nanosatélites capaces

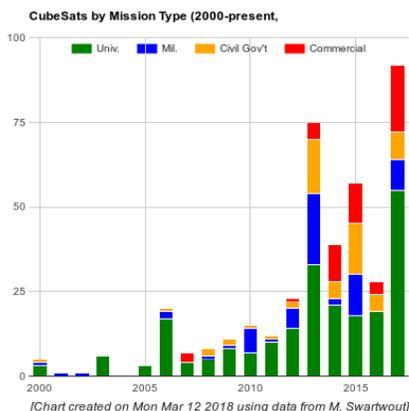


Figura 1: Lanzamientos de CubeSat hasta el 12/03/2018. De: CubeSat Database [5].

de alojar cargas útiles científicas, privadas y gubernamentales [1]. Además, tal como se observa en la Fig. 1, el crecimiento ha sido prácticamente exponencial y principalmente impulsado por universidades. Para acceder a información más detallada sobre distintos proyectos CubeSat se recomienda al lector los trabajos [2, 3, 4].

Usualmente, los desarrolladores de los CubeSat poseen su propia estación terrena para el control de operaciones, pero, debido a la baja órbita terrestre en la que éstos se encuentran, la ventana de tiempo disponible para establecer la comunicación es reducida. Si a esto le sumamos la baja tasa de transferencia que permiten las frecuencias empleadas, principalmente en VHF y UHF, las dificultades a las que se enfrentan los operadores, al intentar aprovechar el satélite en su máxima capacidad, son significativas. Es por ello, que en la actualidad se han desarrollado distintos estándares y proyectos de colaboración internacional de estaciones terrenas interconectadas capaces de programar la adquisición de datos de los distintos satélites no geoestacionarios en su paso alrededor de la superficie terrestre [6, 7].

Desde el punto de vista de la utilización de las imágenes recolectadas desde los nanosatélites, es de destacar que las mismas nos brindan valiosa información de nuestra región. Esto es, si bien las primeras misiones CubeSat apenas



Figura 2: Deforestación en los Andes Bolivianos: 15/06/2016 (izq), 01/01/2017 (der). De: Planet Labs [8].

eran capaces de realizar capturas de baja resolución espacial (100 a 200 m), con el tiempo se han logrado considerables mejoras de resolución (5 a 10 m) e incluso la toma de imágenes multiespectrales, lo que permite un mejor análisis estratégico de los datos recolectados. A modo de ejemplo, en la Fig. 2, se observa el avance de la deforestación de bosques nativos en los Andes Bolivianos para la posterior explotación de la caña de azúcar.

En base a lo antes mencionado, el presente proyecto plantea el desarrollo de un sistema de posicionamiento automático de múltiples antenas que reciban señales (UHF, VHF y S-Band) provenientes de satélites no geoestacionarios con el objetivo de la utilización estratégica de dicha información. Esto es, utilizar las imágenes adquiridas en el monitoreo de la salud de cultivos, la cuantificación de superficies inundadas y el análisis climático de nuestro entorno. Además, se plantea la participación en los distintos proyectos de colaboración internacional existentes.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto consta de los siguientes ejes de investigación y desarrollo:

- Análisis de las necesidades de hardware y software requeridas para el desarrollo del sistema de posicionamiento de antenas y del sistema de comunicación.
- Estudio de la disponibilidad y oferta local de los dispositivos a integrar en el sistema.

- Evaluación de la estrategia de procesamiento digital de imágenes a emplear en base a la necesidad propuesta (monitoreo de la salud de cultivos, cuantificación de superficies inundadas y análisis climático de nuestro entorno).
- Vinculación con los proyectos de colaboración internacional de estaciones terrenas.
- Desarrollo del software y hardware necesarios para el sistema de control de posición de antenas y el sistema de comunicación.
- Generación de documentación adecuada para la discusión interna de resultados.
- Divulgación de los resultados parciales y finales de la investigación.

3. OBJETIVOS

Objetivo General:

Desarrollar un sistema de control automático de posicionamiento de antenas para adquirir señales provenientes de satélites no geoestacionarios, las cuales nos permitan el análisis estratégico de variables de nuestro entorno.

Objetivos específicos:

- Evaluar la eficiencia del sistema de control automático de posicionamiento de antenas y del sistema de comunicación.
- Estimar parámetros de interés estratégico para nuestra región a partir del procesamiento digital de las imágenes recolectadas de los CubeSat.
- Divulgar los resultados y capacidades de la investigación.
- Establecer las bases necesarias para iniciar distintos proyectos de investigación en áreas afines a las tecnologías satelitales.

4. RESULTADOS ESPERADOS

Siendo que el proyecto se encuentra en su fase inicial durante la ejecución del mismo se pretenden alcanzar los siguientes resultados:

- Desarrollo de una estación terrena de posicionamiento automático para la adquisición de señales provenientes de satélites no geoestacionarios.
- Adquisición de imágenes actualizadas de nuestra región.
- Desarrollo de la estrategia de procesamiento de las imágenes recibidas de acuerdo a cada necesidad propuesta.
- Colaboración con los proyectos internacionales existentes.
- Divulgación en el medio de las capacidades del proyecto y sus resultados.
- Formación de recursos humanos capaces de continuar proyectos afines a esta línea de investigación.

5. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo abocado al proyecto se encuentra integrado al Grupo de Investigación y Desarrollo Electrónico (GIDE), el cual está conformado por su director Mg. Ing. Gastón Peretti, su co-director Ing. Sergio Felissia, los investigadores de apoyo Ing. Emanuel Bernardi e Ing. Hugo Pipino y los alumnos avanzados Facundo Busano, Marco Miretti y Emmanuel Dovis. Es de destacar que la presente investigación está estipulada como parte de la tesina de grado de al menos dos de los alumnos involucrados.

6. BIBLIOGRAFÍA

- [1] California Polytechnic State University. *CubeSat Design Specification Revision 13*. Inf. Téc. 2014.
- [2] Egil Eide y Jorgen Iltad. «NCUBE-1 the First Norwegian Cubesat Student

- Satellite». En: *16th ESA Symposium on European Rocket and Ballon Programmes and Related Research* (jun. De 2003).
- [3] Roger Birkeland y Odd Gutteberg. «Overview of the NUTS CubeSat Project». En: *International Academy of Astronautics - Cubesat Workshop* (2009).
- [4] Matthew Patrick Schroer. «NPS-SCAT: a CubeSat Communications System Design, Test, and Integration». Tesis doct. 2009
- [5] Michael Swartwout. CubeSat Database. URL:<https://sites.google.com/a/slu.edu/swartwout/home/cubesat-database> (visitado 12-03-2018).
- [6] Daniel J White y col. «SatNOGS: Satellite Networked Open Ground Station». En: *Engineering Faculty Publications* (oct. de 2015). URL : <https://satnogs.org/>.
- [7] Graham Shirville y Bryan Klofas. «Genso: a Global Ground Station Network». En: *AMSAT Symposium* (2007).
- [8] Planet Labs. *Deforestación por la caña de azúcar en los Andes Bolivianos*. URL: <https://www.planet.com/gallery/sugarcane-deforestation/> (visitado 14-03-2018).

SISTEMA DE COMUNICACIÓN PARA LA IMPLEMENTACIÓN DE REDES INALÁMBRICAS DE SENSORES

Héctor Hugo Mazzeo hmgvm@yahoo.com; José A. Rapallini josrap@gmail.com;
Omar E. Rodríguez rodriguezomarlp@gmail.com; Marcelo Zabaljauregui mzabaljauregui@gmail.com;
Universidad Tecnológica Nacional – Facultad Regional La Plata
La Plata, Buenos Aires, Argentina

RESUMEN

El objetivo de este proyecto es el estudio de los protocolos de comunicación de las redes de sensores inalámbricos y sus aplicaciones en diversos ámbitos. Si bien en general estas redes utilizan Zigbee como protocolo de comunicación inalámbrico y módulos Xbee como nodos de red, buscamos investigar otras opciones más económicas para su implementación. La idea es utilizar la experiencia adquirida en el manejo de la plataforma Arduino con sus shields o módulos inalámbricos para implementar la comunicación a través de la red. A pesar de tener algunas limitaciones con respecto a la tecnología Zigbee, éstas pueden ser soslayadas en algunas aplicaciones prácticas donde el consumo de energía no sea un factor crítico o la cantidad de sensores sea baja. También habrá que tener en cuenta las distancias entre nodos, ya que uno de los fuertes de Zigbee es el alcance.

Palabras clave: sensores inalámbricos, nodos, motes, red WiFi, gateway.

CONTEXTO

El proyecto está enmarcado dentro del PID SIUTNLP0005017, elaborado y en ejecución por parte del grupo de Codiseño hardware/software para Aplicaciones de Tiempo Real (CODAPLI), cuyo laboratorio pertenece al Departamento de Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional La Plata, institución que financia íntegramente el proyecto, facilitando asimismo las instalaciones y equipamiento para su realización. Las tareas principales se orientan a la investigación y desarrollo de proyectos relacionados con los sistemas de tiempo real:

sistemas centralizados y distribuidos basados en computadoras personales, control y adquisición de datos, sistemas embebidos basados en placas con microcontrolador, autómatas programables, robótica, etc.

1. INTRODUCCIÓN

Las redes de sensores inalámbricos (Wireless Sensor Networks, WSN) son redes inalámbricas formadas por dispositivos autónomos distribuidos espacialmente que utilizan sensores para monitorear condiciones físicas o ambientales. Un sistema WSN incorpora un gateway que provee conectividad inalámbrica con los nodos distribuidos haciendo de interfaz con una estación base recolectora de datos, normalmente conectada a una red TCP/IP (Fig. 1). El protocolo inalámbrico seleccionado depende de los requerimientos de la aplicación. Algunos de los estándares disponibles incluyen radios de 2.4 GHz basados en los estándares IEEE 802.15.4 o IEEE 802.11 (Wi-Fi) o radios propietarios, los cuales son regularmente de 900 Mhz [1].

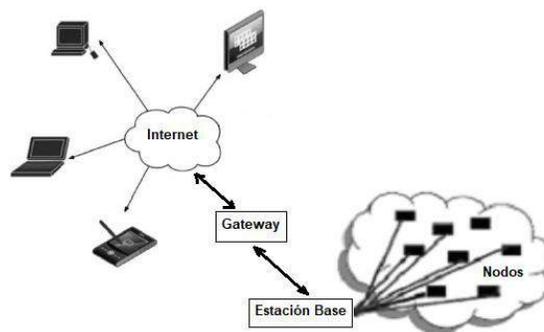


Fig. 1. Esquema básico de una WSN

El desarrollo de WSNs fue inspirado por aplicaciones militares, especialmente en la

vigilancia de las zonas de conflicto. Se componen de pequeños nodos de sensores equipados con interfaces de radio y están distribuidas sobre una región geográfica. La tarea de cada sensor es realizar mediciones y enviar datos a un nodo coordinador. El avance en la tecnología ha impulsado la implementación de redes de sensores compuestas por nodos de bajo costo (llamados comúnmente "motes"), los cuales están formados por el sensor propiamente dicho, la interfase para realizar las comunicaciones y, generalmente, un microprocesador. Estos nodos están diseñados para adquirir información del entorno, procesarla y transmitir los datos pertinentes a la estación base (de mucho mayor poder computacional que los nodos sensores). El campo de aplicación para esta tecnología es muy amplio. Algunos usos actuales son: control de procesos, alarmas, respuesta a emergencias, redes de transporte, investigación biológica, aplicaciones medicinales y estudios meteorológicos.

Existen sensores de todas clases, es posible medir aceleración, temperatura, movimientos sísmicos, posición global, intensidad de luz, sonido, campos magnéticos, etc. La interconexión de los sensores se presenta como uno de los mayores problemas a la hora de implementar el sistema. Las limitaciones físicas y económicas imponen restricciones a la forma en la que se llevará a cabo. Anteriormente, las redes de sensores se interconectaban mediante cables (pares trenzados) que agregaban costo, complejidad de implementación y dificultad en la detección de fallas. Con la estandarización de las comunicaciones inalámbricas y la serialización de fabricación de dichos productos, ésta pasó a ser una de las formas preferenciales para interconectar sensores. Actualmente, los sensores inalámbricos son de amplio uso en aplicaciones distribuidas o que implican el seguimiento de objetivos móviles. La administración del consumo de energía para este tipo de sensores constituye un factor crítico en el diseño de dichos sistemas debido a que los mismos están generalmente alimentados mediante baterías.

Además del problema del consumo, existe la dificultad de su reemplazo. La energía es utilizada por los nodos sensores en tres funciones principales: sensado, procesamiento y comunicación. Los dispositivos modernos utilizan técnicas avanzadas de ahorro de energía como el apagado de periféricos en desuso (stand-by) o reducción de la frecuencia de operación. Se ha demostrado que el proceso que requiere mayor energía es la transferencia de datos, por lo que a veces se apunta a reducir la cantidad de comunicaciones a costa de mayor tiempo de procesamiento [2], [3].

Las redes de sensores se caracterizan por ser redes desatendidas (sin intervención humana), con alta probabilidad de fallo (en los nodos, en la topología), habitualmente construidas *ad hoc* para resolver un problema muy concreto (es decir, para ejecutar una única aplicación). Las comunicaciones inalámbricas son una solución eficaz y confiable en la automatización del hogar y la oficina, en tal sentido, varios medios de transmisión de señales podrían utilizarse, incluyendo la luz y el ultrasonido, pero haciendo una comparación entre ellos con respecto al posible tráfico de datos, los precios y el área de cobertura, terminan finalmente siendo más atractivos los enlaces de radio frecuencia (RF) [4].

Se han creado aplicaciones WSN para diferentes áreas incluyendo cuidado de la salud, servicios básicos y monitoreo remoto. En el cuidado de la salud, los dispositivos inalámbricos vuelven menos invasivo el monitoreo a pacientes. Para servicios básicos como electricidad, alumbrado público y agua corriente, los sensores inalámbricos ofrecen un método de bajo costo para un sistema de recolección de datos que ayude a reducir el uso de energía y mejor manejo de recursos. El monitoreo remoto cubre un amplio rango de aplicaciones donde los sistemas inalámbricos pueden complementar sistemas cableados reduciendo costos en cables y permitiendo nuevos tipos de aplicaciones de medición.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación se orientan principalmente en la búsqueda de alternativas más económicas y simples tanto en la comunicación inalámbrica entre los nodos como en su diseño y desarrollo. Una opción es el uso de la tecnología Arduino con sus placas de bajo costo y las posibilidades de interconexión vía WiFi a través de shields o módulos inalámbricos frente a las más costosas placas XBee que utilizan el standard Zigbee para la comunicación inalámbrica (Fig. 2). En aquellos sistemas que no requieran una gran cantidad de nodos ni el mayor consumo de energía sea un problema, puede resultar en una solución aceptable y de bajo costo [5], [6].



Fig. 2. Prototipo de sistema transmisor-receptor con módulos XBee desarrollado en el laboratorio

3. RESULTADOS OBTENIDOS/ESPERADOS

En CODAPLI hemos investigado gran parte de esta temática, desarrollado varios proyectos e informes técnicos los cuales, si bien no han desembocado en un sistema integral basado en el control por redes de sensores inalámbricos, incluyen elementos o partes de estos sistemas: sensores, sistemas embebidos, comunicaciones inalámbricas entre placas microcontroladas, sensado remoto y automatización de dispositivos.

Se ha experimentado y realizado prototipos de adquirentes de datos y comunicaciones inalámbricas utilizando diferentes tecnologías y plataformas: Arduino, Intel-Galileo, EDUCIAA (Fig. 3). Esto nos permite tener una cierta experiencia a la hora de evaluar diferentes implementaciones, especialmente en el tema de la comunicación inalámbrica sobre la red de sensores, tema principal del proyecto.



Fig. 3. Aplicación móvil para toma de datos y control de dispositivos físicos mediante combinación de placas Arduino-Intel Galileo.

A modo de ejemplo mencionamos algunos de los principales trabajos realizados por el grupo, varios de los cuales fueron presentados y expuestos en diversas jornadas y congresos afines a la temática:

- Plataforma remota para prácticas de laboratorio (WICC 2017).
- Sistema de Monitoreo Meteorológico utilizando la metodología de codiseño hardware/software (44 JAIIO 2015).
- Estación de medición para análisis y control de parámetros ambientales (46 JAIIO- CAI9 2017).
- Prototipo de Estación Meteorológica Automática utilizando EDUCIAA-NXP como plataforma (CNEISI 2016; Fig. 4).
- Comunicación por protocolo Zigbee, prestaciones, aplicación al mundo de domótica (informe técnico).

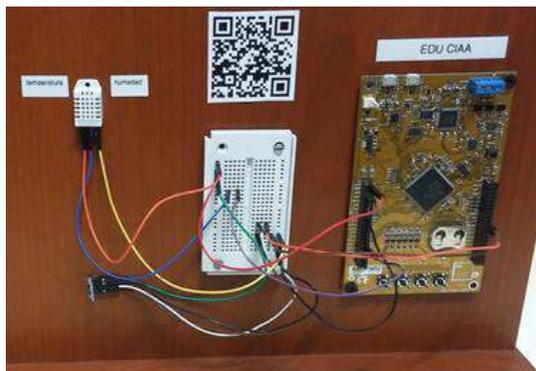


Fig. 4. Prototipo de estación meteorológica utilizando una placa EDU-CIAA para la adquisición de datos.

Basándonos en la experiencia obtenida en estos años de investigación, se espera desarrollar aplicaciones orientadas a la utilización de sistemas ciberfísicos (mecanismos controlados o monitoreados por algoritmos basados en computadoras y estrechamente integrados con Internet y sus usuarios) mediante el control y la adquisición de datos basados en la comunicación a través de redes inalámbricas de sensores. Se pretende que estas aplicaciones puedan interactuar con los sistemas físicos que lo rodean pudiendo ser monitorizadas en forma local y remota a través de los medios tradicionales como Internet, utilizando computadoras, tablets, PDA, etc. Los campos de aplicación son variados y extensos, siendo la idea principal poder orientarlo principalmente al ámbito industrial, pero no descartando otros como la salud, educación, agricultura, meteorología, Internet de las cosas, etc.

La propuesta es avanzar en distintos ejes de investigación, aprovechando la experiencia y los conocimientos adquiridos por los integrantes del grupo. Por ello se deberá encarar el proyecto a través de distintos frentes, abordados por uno o más integrantes, quienes se enfocarán en un tema particular, pero a su vez integrados en un todo con el resto. Podemos entonces identificar los principales temas o ramas sobre las cuales enfocar la investigación:

- estudio de distintos sensores;

- conexasión y comunicación mediante las interfaces y protocolos más utilizados;
- estudio de las comunicaciones inalámbricas entre motes de la red.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto está orientado a la formación de recursos humanos particularmente escasos en nuestro país, en particular para esta temática. Se busca también lograr la transferencia de conocimientos a los alumnos que requieran una formación adicional a la recibida en las cátedras de especialidades afines al tema. En el plantel básico del grupo se incorporan periódicamente graduados noveles y estudiantes del último año. La propuesta a mediano y largo plazo es el dictado de seminarios y cursos de grado y posgrado para aumentar la capacitación en estos temas.

El equipo de trabajo que participa actualmente en el desarrollo de éste y otros proyectos de investigación relacionados a la temática de los sistemas de tiempo real está conformado por dos Ingenieros en Electrónica (director y codirector del grupo) y dos Ingenieros en Sistemas que también se desempeñan como docentes en el área de Sistemas de Información de la UTN-FRLP, uno de los cuales se encuentra realizando su tesis de postgrado. El grupo se completa con cuatro becarios, siendo tres de ellos alumnos y un graduado.

Colaboran también en forma *ad honorem* alumnos del último año de la carrera de Ingeniería en Sistemas realizando tesinas de grado que cubren aspectos relacionados con esta temática en particular y otros aspectos relacionados con los sistemas de tiempo real.

5. BIBLIOGRAFÍA

- [1] William Stalling. Comunicaciones y Redes de Computadoras, Pearson 7^a edición, 2016.
- [2] Chaudhari, Dharavath. "Study of Smart Sensors and their Applications".

- Proceedings on International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), Vol. 3, Issue 1, January 2014, ISSN (Online): 2278-1021.
- [3] IEC, 2014 “Internet of Things: Wireless Sensor Networks” Proceedings on International Electrotechnical Commission (IEC), ISBN 978-2-8322-1834-1.
- [4] Lutokhin. “Applications of Wireless Sensor Networks in Next Generation Networks”. Proceedings of International Telecommunication Union (ITU), 28 February 2014..
- [5] Shahin Farahani. Zigbee Wireless Networks and Transceivers. Elsevier,2008.
- [6] ZigBee Alliance. ZigBee Standards-<http://www.zigbee.org/Standards/Overview.aspx>, 2011.

Sistemas Embebidos y de Tiempo Real con Planificación Heterogénea

José M. Urriza, Mariano A. Ferrari, Javier D. Orozco, Francisco E. Páez,
Gabriela Olguín, Fernando Tidona, Rodrigo Tolosa

Depto. de Informática – Facultad de Ingeniería – Universidad Nacional de la
Patagonia San Juan Bosco – Sede Puerto Madryn
Bvld. Brown 3051, Puerto Madryn, Chubut
josemurriza@gmail.com

Resumen

Este proyecto plantea, como objetivo general, el desarrollo de técnicas de modelado, diseño, análisis, optimización y prueba de *Sistemas Embebidos* con requerimientos de *Tiempo Real*, específicamente del tipo Heterogéneo, sobre diferentes plataformas hardware y software. El objetivo es lograr un balance apropiado de prestaciones, rendimiento y eficiencia en el desarrollo de nuevos sistemas.

Palabras clave: Sistemas Embebidos, Sistemas de Tiempo Real, Planificación, Sistemas Operativos de Tiempo Real

Contexto

El proyecto “*Sistemas Embebidos y de Tiempo Real con Planificación Heterogénea*” se inserta en las líneas de investigación en Sistemas de Tiempo Real (*STR*), Sistemas Embebidos (*SE*) y de propósito dedicado, y Sistemas Operativos de Tiempo Real (*SOTR*). Particularmente, se busca resolver problemas concretos de implementación y

planificación sobre plataformas de desarrollo reales disponibles en el mercado, utilizadas tanto en la industria como en el campo educacional y de investigación, como por ejemplo Arduino, Arduino Due, mbed, EDUCIAA, Raspberry Pi, PCduino, FRDM K64F, STM32F7, etc. El grupo de investigación en STR de la Sede Puerto Madryn de la UNPSJB¹ posee varios ejemplares de cada una de estas plataformas. El proyecto será acreditado por la Secretaría de Ciencia y Técnica de la UNPSJB, la cual también es responsable del financiamiento del mismo. Además, se espera una estrecha cooperación como asesor, con el Dr. Javier Orozco, del Departamento de Ingeniería Eléctrica y Computadoras de la UNS.

Introducción

Actualmente, todo país industrializado, debe contar con profesionales capacitados en *SE* y *STR*, ya que los mismos se han expandido más allá de las industrias clásicas como la militar, la aeroespacial, la médica, la de control de procesos, etc.,

¹ <http://www.rtsg.unp.edu.ar/>

para formar parte de la gran mayoría de los dispositivos electrónicos que se emplean cotidianamente como los teléfonos celulares, los automóviles, los electrodomésticos, etc. Mundialmente los *STR* son una línea de investigación recurrente en numerosos centros de investigación, tanto en universidades como en las industrias, con el objetivo de proveer de herramientas de diseño y soluciones que permitan maximizar la eficiencia del uso de los recursos computacionales, minimizando el costo de los productos y maximizando el ciclo de vida de sus productos. Por ejemplo, un diseño apropiado permite cumplir los requerimientos del sistema, brindando además ejecución de heterogéneas aplicaciones (que se llamará de aquí en adelante conjuntos de tareas heterogéneas). Por otro lado, es importante mantener un ahorro de energía, poseer tolerancia a las fallas, entre otras diversas funciones. Es importante, además, reducir la complejidad de implementación, tanto desde el punto de vista del hardware como del software, sin sacrificar robustez ni eficiencia. Para ello, nuevas herramientas de desarrollo, que permitan validar especificaciones, son necesarias.

Dentro de las principales motivaciones para el desarrollo de este proyecto, se pueden nombrar los siguientes puntos:

- Los *SE* son utilizados constantemente en la industria. Estos sistemas cuentan con restricciones temporales, y, por lo tanto, generalmente, emplean un *SOTR*. Sin embargo, en muchos casos la implementación no suele ser eficiente.
- Además, en los *STR*, es complejo alcanzar altos niveles de eficiencia, garantizando al mismo tiempo las constricciones temporales. Esto es particularmente problemático en *SE*, donde usualmente se requiere un

compromiso entre el costo de la plataforma hardware, y los objetivos del sistema. Por ejemplo, sobredimensionar la potencia de cálculo del microprocesador genera un incremento en el consumo de energía, lo que conlleva un mayor tamaño de la batería, menor autonomía, etc. Existen numerosos desarrollos teóricos, sin implementación o prueba práctica, que podrían resolver estas cuestiones para ciertas clases de aplicaciones.

- Es necesario adecuar las herramientas existentes, y desarrollar nuevas, para abordar estas problemáticas. Además, es necesario poder evaluar el rendimiento de estos desarrollos en casos concretos, lo que implica generar y probar prototipos sobre plataformas de desarrollo, tales como Arduino, mbed, Raspberry, etc.
- Existen también motivaciones socioeconómicas, como la reducción de los costos de producción, menor consumo de energía y contaminación ambiental, mejores prestaciones y ciclo de vida del producto, etc.
- Regionalmente, es una motivación generar nuevos recursos humanos, con alta capacidad y competitividad, que contribuyan al desarrollo del sector en la región. Para esto, el proyecto incluye alumnos avanzados, a los que se capacitará en la disciplina.

El estado del arte de este proyecto es sumamente extenso, y cubre las disciplinas de *STR*, *SE*, Sistemas de Control, entre otras. Cabe resaltar que la disciplina de *STR* puede considerarse como básica, y por lo tanto engloba a las otras cuando las mismas cuentan con requerimientos temporales. En particular, dentro de los métodos de evaluación de

planificabilidad, que permiten garantizar las constricciones temporales del *STR* al momento de su diseño, los trabajos más relevantes son [1, 2, 3, 4, 5]. Para la planificación heterogénea, los trabajos más importantes son [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]. En cuanto al ahorro de energía, los trabajos más relevantes son [18, 19, 20] para planificación mediante prioridades fijas, [21, 22] para prioridades dinámicas y [23] para el caso de prioridades mixtas. Trabajos que hacen uso de Slack Stealing (SS) son [24, 25]. Por otro lado, en sistemas de control, la planificación con Jitter es un importante problema, abordado en trabajos como [26, 27, 28].

Líneas de Investigación, Desarrollo e Innovación

Como ya se comentó previamente, el presente proyecto se inserta en las líneas de investigación en Sistemas de Tiempo Real (*STR*), Sistemas Embebidos (*SE*) y de propósito dedicado, y Sistemas Operativos de Tiempo Real (*SOTR*).

Se realizará un trabajo de revisión de la biografía existente relevante, para estudiar los modelos y técnicas ya existentes, e interiorizarse en sus aplicaciones y limitaciones. A su vez, se estudiarán las soluciones comerciales y/o gratuitas más relevantes, como los *SOTR* y los diversos kits de desarrollo disponibles.

Las líneas de investigación se centran en las áreas de planificación heterogénea, evaluación de planificabilidad, modelos de ahorro de energía, etc., desarrollando nuevos métodos y técnicas, buscando principalmente mejorar la eficiencia en el uso de los recursos, y facilitar la implementación práctica de los resultados teóricos.

Remarcar que los integrantes de la unidad académica poseen varios trabajos publicados en las anteriores áreas.

Resultados y Objetivos

El objetivo general de este proyecto, es el desarrollo de técnicas de modelado, diseño, análisis, optimización y testeado de *SE* con requerimientos de Tiempo Real Heterogéneos, en múltiples plataformas hardware/software, que permitan lograr un apropiado balance entre costos, prestaciones, rendimiento y eficiencia.

Los objetivos específicos pueden clasificarse como sigue:

- Estudiar la implementación de los *SOTR* más utilizados, y en particular como los mismos administran los recursos computacionales.
- Determinar la factibilidad de aplicación de los modelos y métodos teóricos, disponibles en la literatura pero sin implementaciones en entornos prácticos. Proponer las mejoras necesarias para lograr su implementación.
- Implementar técnicas y métodos propuestos en la literatura en plataformas de desarrollo concretas, para estudiar su comportamiento real. Para esto, se realizarán modificaciones sobre los núcleos de los *SOTR*, o se empleará una tarea planificadora, que implemente el método sin realizar alteraciones en el *SOTR*.
- Publicar los resultados de las investigaciones realizadas, fomentando e incentivando la difusión de las tareas de investigación de la UNPSJB, y consecuentemente aportando a la formación de recursos humanos.

En cuanto a los resultados concretos del proyecto, los mismos aún no pueden ser evaluados dado que la fecha de inicio del mismo es Octubre del presente año. Se espera que los resultados de las investigaciones generen desarrollos que cumplan con los objetivos planteados, en forma de nuevos diseños, algoritmos y métodos, implementados sobre plataformas hardware/software y SOTR concretos, en forma de prototipos. Esto permitirá evaluar el costo computacional real de los métodos teóricos.

Formación de Recursos Humanos

En este proyecto de investigación trabajarán en un principio 7 (siete) personas, en las cuales 6 (seis) son docentes y un alumno. De las mismas, 3 cuentan con el grado de Doctor, y una está en proceso de obtener el título de Doctor. Además, se espera sumar otros alumnos avanzados de la carrera Licenciatura en Informática de la Sede Puerto Madryn de la UNPSJB, con el objetivo de que puedan iniciar sus Tesinas de Grado en el marco del proyecto.

Referencias

- [1] J. M. Urriza, F. E. Páez, M. Ferrari, R. Cayssials, and J. D. Orozco, "A New RM/DM Low Cost Schedulability Test," in *Eight Argentine Symposium and Conference on Embedded Systems*, Buenos Aires, 2017, pp. 13-18.
- [2] C. L. Liu and J. W. Layland, "Scheduling Algorithms for Multiprogramming in a Hard Real-Time Environment," *Journal of the ACM*, vol. 20, pp. 46-61, 1973.
- [3] M. Sjödin and H. Hansson, "Improved Response-Time Analysis Calculations," in *IEEE 19th Real-Time Systems Symp.*, 1998, pp. 399-409.
- [4] M. Joseph and P. Pandya, "Finding Response Times in Real-Time System," *The Computer Journal (British Computer Society)*, vol. 29, pp. 390-395, 1986.
- [5] E. Bini, G. Buttazzo, and G. Buttazzo, "A Hyperbolic Bound for the Rate Monotonic Algorithm," *IEEE Transactions on Computer*, vol. 52, pp. 933-942, 2003.
- [6] R. M. Santos, J. M. Urriza, J. Santos, and J. D. Orozco, "New methods for redistributing slack time: applications and comparative evaluations," *The Journal of Systems & Software*, vol. 70-2, pp. 115-128, 2004.
- [7] J. M. Urriza, R. Cayssials, and J. D. Orozco, "A Fast Slack Stealing Method for embedded Real-Time Systems," Dep. de Ing. Eléctrica y Computadoras, Universidad Nacional del Sur, Argentina., Bahía Blanca, Internal Report May 31 2005.
- [8] J. M. Urriza, F. E. Páez, R. Cayssials, J. D. Orozco, and L. Schorb, "Low Cost Slack Stealing Method for RM/DM," *International Review in Computers and Software (IRECOS)*, vol. 5, pp. 660-667, 2010.
- [9] F. Páez, J. M. Urriza, R. Cayssials, and J. D. Orozco, "Métodos de Slack Stealing en FreeRTOS," in *44 Jornadas Argentinas de Informática (JAIIO)*, Rosario, 2015.
- [10] F. E. Páez, J. M. Urriza, R. Cayssials, and J. D. Orozco, "FreeRTOS user mode scheduler for mixed critical systems," in *Embedded Systems (CASE), 2015 Sixth Argentine Conference on*, 2015, pp. 37-42.
- [11] R. I. Davis, K. W. Tindell, and A. Burns, "Scheduling Slack Time in Fixed-Priority Preemptive Systems," *Proceedings of the Real Time System Symposium*, pp. 222-231, 1993.
- [12] R. I. Davis, "Approximate Slack Stealing Algorithms for Fixed Priority Pre-emptive Systems," Real-Time Systems Research Group, University of York, York, England, Internal Report 1994.
- [13] R. Davis and A. Wellings, "Dual priority scheduling," in *Real-Time Systems Symposium, 1995. Proceedings., 16th IEEE*, 1995, pp. 100-109.
- [14] J. P. Lehoczky and S. Ramos-Thuel, "An Optimal Algorithm for Scheduling Soft-Aperiodic Tasks in Fixed-Priority Preemptive Systems," in *IEEE Real-Time*

- Systems Symposium*, Phoenix, Arizona, EUA, 1992, pp. 110-123.
- [15] S. Ramos-Thuel and J. P. Lehoczky, "On-Line Scheduling of Hard Deadline Aperiodic Tasks in Fixed-Priority Systems," in *Real-Time Systems Symposium*, 1993, pp. 160-171.
- [16] S. Ramos-Thuel and J. P. Lehoczky, "Algorithms for Scheduling Hard Aperiodic Tasks in Fixed-Priority Systems using Slack Stealing," in *Real-Time Systems Symposium*, 1994, pp. 22-33.
- [17] T.-S. Tia, J. W. Liu, and M. Shankar, "Aperiodic Request Scheduling in Fixed-Priority Preemptive Systems," Department of Computer Science, University of Illinois, Internal Report UIUCDCS-R-94-1859, 1994.
- [18] B. Novelli, J. C. B. Leite, J. M. Urriza, and J. D. Orozco, "Regulagem Dinâmica de Voltagem em Sistemas de Tempo Real," in *XXXII Seminário Integrado de Software e Hardware (SBC 2005 SEMISH)*, Unisinos -Sao Leopoldo, Brazil, 2005.
- [19] S. Saewong and R. Rajkumar, "Practical Voltage-Scaling for Fixed-Priority RT-Systems," in *9th IEEE Real-Time and Embedded Technology and Applications Symposium*, Toronto, Canada, 2003, pp. 106-115.
- [20] Y. Shin and K. Choi, "Power conscious fixed priority scheduling for hard real-time systems," in *36th Design Automation Conference*, 1999, pp. 134-139.
- [21] P. Pillai and K. G. Shin, "Real-Time Dynamic Voltage Scaling for Low-Power Embedded Operating Systems," in *18th Symposium on Operating Systems Principles*, Banff, Alberta, Canada, 2001, pp. 89-102.
- [22] W. Kim, J. Kim, and S. L. Min, "A Dynamic Voltage Scaling Algorithm for Dynamic-Priority Hard Real-Time Systems Using Slack Time Analysis," in *Conference on Design, Automation and Test in Europe*, Washington, DC, EUA, 2002, pp. 788-794.
- [23] M. A. Moncusí, A. Arenas, and J. Labarta, "Improving Energy Saving in Hard Real Time Systems via a Modified Dual Priority Scheduling," vol. 29, pp. 19-24, 2001.
- [24] J. M. Urriza, R. Cayssials, J. D. Orozco, and J. C. B. Leite, "Modelo de Tareas para recuperacion de Slack para Aplicaciones en Sistemas Enbebidos con DVS," Dep. de Ing. Eléctrica y Computadoras, Universidad Nacional del Sur, Argentina., Bahía Blanca, Reporte Interno 22 de Mayo 2005.
- [25] J. M. Urriza, J. D. Orozco, C. Buckle, and R. Cayssials, "Ahorro de Energía en Dispositivos con un SO de Tiempo Real que planifican en RM o DM," in *Encuentro Chileno de Computacion*, Santiago, Chile, 2009.
- [26] K. W. Tindell, "Fixed Priority Scheduling of Hard Real-Time Systems," Doctor of Philosophy, Department of Computer Science, University of York, 1993.
- [27] N. C. Audsley, A. Burns, M. F. Richardson, K. Tindell, and A. J. Wellings, "Applying New Scheduling Theory to Static Priority Preemptive Scheduling," *Software Engineering Journal*, vol. 8, pp. 284-292, 1993.
- [28] R. Rajkumar, "Real-time synchronization protocols for shared memory multiprocessors," in *Distributed Computing Systems, 1990. Proceedings., 10th International Conference on*, 1990, pp. 116-123.

Visión artificial aplicada en Agricultura de Precisión.

Claudia Russo^{1,2}, Hugo Ramón^{1,2}, Sandra Serafino¹, Benjamin Cicerchia^{1,4}, Mónica Sarobe¹, Agustín Balmer^{1,3}, Álvarez Eduardo¹, Pablo Luengo¹, Gustavo Useglio¹, Martín Faroppa^{1,3}

Instituto de Investigación y Transferencia en Tecnología (ITT)⁵
Comisión de Investigaciones Científicas (CIC)
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)
Sarmiento y Newbery, Junín (B) – TE: (0236) 4477050

{claudia.russo, hugo.ramon, sandra.serafino, lucas.cicerchia, monica.sarobe, agustin.balmer, eduardo.alvarez, pablo.luengo, gustavo.useglio, martin.faroppa}@itt.unnoba.edu.ar

Resumen

La visión artificial representa hoy un área de gran utilidad e interés para los investigadores más allá de que sus técnicas se remiten a más de 3 décadas de desarrollos. Esto se debe a la expansión tecnológica que ha permitido una generación de nuevos procesadores, nuevos sensores, y nuevas capacidades de almacenamiento que han sostenido el desarrollo de nuevas técnicas de inteligencia artificial y en particular de visión. El nivel de automatismos se ha incrementado exponencialmente en los últimos años. Las aplicaciones son de lo más variadas, desde la utilización de navegadores de todo tipo de vehículos (p.e plataformas robóticas), detección de características en la observación de objetos (p.e. líneas de producción, líneas de cultivos), clustering y clasificación automatizada de objetos (p.e clasificación de granos), tracking de objetos fijos o móviles (p.e. seguimiento de personas).

En la UNNOBA se está trabajando en la aplicación combinada de diferentes desarrollos tecnológicos aplicados a Agricultura de Precisión, muchos basados en el uso de imágenes y técnicas de visión. Entre ellos el desarrollo de procesos de navegación automatizada de ensayos a campo y el relevamiento de datos, indispensables para el ajuste adecuado de los procesos de mejoramiento genético de semillas.

Palabras clave: Visión artificial, Robótica, Imágenes, Agricultura de Precisión.

Contexto

Esta línea de investigación forma parte del proyecto “Informática y Tecnologías Emergentes” aprobado por la Secretaría de Investigación, Desarrollo y Transferencia de la UNNOBA en el marco de la convocatoria a Subsidios de Investigación BIANUALES (SIB 2017). A su

¹ Docente Investigador en el Instituto de Investigación y Transferencia en Tecnología (ITT) / Escuela de Tecnología / UNNOBA

² Investigador Asociado Adjunto sin director a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

³ Becario PROMINF / ITT / Escuela de Tecnología

⁴ Becario Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

⁵ Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

vez se enmarca en el contexto de planes de trabajo aprobados por la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires y por la Secretaría de Investigación de la UNNOBA en el marco de la convocatoria “Becas de Estudio Cofinanciadas 2015 CIC Universidades del interior bonaerense”.

A esto se suma el proyecto “Fenotipado de Alta Capacidad con Relevamiento de Datos en Campo” el que se encuentra actualmente en desarrollo. Este es un trabajo conjunto entre la Secretaría de Investigación, Desarrollo y Transferencia de la UNNOBA y el INTA, Estación Experimental Pergamino. El mismo se desarrolla en el Instituto de Investigación en Tecnologías y Transferencia (ITT) dependiente de la mencionada Secretaría, y se trabaja en conjunto con la Escuela de Tecnología de la UNNOBA. Sobre la base este proyecto se plantean además avances y mejoras orientadas al desarrollo de plataformas multipropósitos y auto navegables mediante visión artificial, que puedan coordinarse con vehículos aéreos utilizando estas mismas técnicas sumadas a geoposicionamiento mediante DGPS.

El equipo está constituido por profesionales de ambas instituciones, UNNOBA e INTA. Por parte de la UNNOBA intervienen docentes e investigadores pertenecientes al ITT, al Departamento de Ciencias Agrarias, Escuela de Ciencias Agrarias así como también, estudiantes de las carreras de Informática de la Escuela de Tecnología de la UNNOBA. Y por parte del INTA, intervienen; Mejoradores Genéticos, Ecofisiólogos e Ingenieros Agrónomos.

Introducción

En los últimos años el avance de la tecnología en la Agricultura de Precisión (AP) ha sido exponencial, no solo aplicada a la mejora del cultivo, sino en las herramientas de trabajo con el fin de lograr más y mejores rindes. Por ejemplo aplicados a maquinaria para el monitoreo de siembra o con un sistema de navegación autónomo en un tractor. A esto se le sumó la utilización de diferentes tipos de plataformas robóticas móviles [1], ya sean éstas aéreas o terrestres. En ese aspecto se encuentran diferentes desarrollos tanto en el ámbito universitario como en la empresa. Entre las universidades se destaca el Centro Australiano de Robots a Campo de la Universidad de Sydney [2] que es una de las pioneras, con plataformas terrestres de sensado y control del cultivo entre otras, al igual que el Instituto de Robótica de la Universidad de Carnegie Mellon [3]. Y en el ámbito privado se destacan, la empresa Bosch con una plataforma terrestre de sensado [4] y la empresa Agco/Fendt [5] con sus plataformas terrestres de sembrado, o la empresa DJI [6] con sus drones para monitoreo general del cultivo o fumigación.

A ello se le suma el avance de la tecnología en las cámaras digitales, que hace de estas una herramienta fundamental a la hora de sensar o extraer información de lo que se desea observar. En particular, en el caso de los sistemas de visión artificial [7], necesarios entre otros, para llevar adelante la navegación automatizada de los vehículos, es importante destacar que los mismos

pueden clasificarse en sistemas directos e indirectos. Los primeros son aquellos que permiten trabajar en tiempo real (como en el caso de los vehículos terrestres)[8][9][10], y los sistemas indirectos son aquellos cuyo procesamiento es posterior a la captura de datos obtenidos mediante imagen y/o video (como en el caso de los UVAs). Esta diferenciación los sitúa generalmente en aplicaciones de diferente utilidad, siendo la utilización de UVAs aplicados a la recolección de información general de zonas de cultivos que abarcan más de un surco o parcela, y la utilización de sistema de visualización directa a la navegación intra/extra surcos y/o caminos y la recolección de datos en general fenotípicos de dichos cultivos [11][13]. En cualquier caso será necesario el procesamiento digital de imágenes captadas por el sistema de visión artificial de cualquiera de los tipos mencionados y toda la problemática que el relevamiento de este tipo de datos genera, como por ejemplo el cálculo efectivo de distancias [14] en los sistemas directos y/o la corrección de iluminación en ambos casos [15].

Líneas de Investigación, Desarrollo e Innovación

Entre las líneas de investigación relacionadas con la temática expuesta, se destacan 3 líneas principales.

A partir de la creación de una plataforma robótica móvil terrestre de navegación de cultivos, se espera dotar a la misma de un sistema de visión artificial que le permita a esta poder navegar el cultivo intra/extra

surcos y/o a través de caminos.

Por otro lado, la plataforma busca recolectar datos biofísicos del cultivo a lo largo de la evolución de sus estados fenológicos, sumado al monitoreo y evaluación de los mismos [12][13]. Para obtener dichos datos la plataforma debe sensar el cultivo mediante la captura de imágenes, a partir de un recorrido establecido en el ensayo, para lo cual debe determinar en qué momento del recorrido se deberán sensar las parcelas. Para determinar dicho instante el sistema debe ser capaz de detectar los diferentes plantines del cultivo, teniendo en cuenta además cuales de ellos ya fueron sensados. A partir de esto es que se dispone anexar a la plataforma la utilización de una cámara de video que permita realizar un tracking de las diferentes parcelas. Es decir un sistema de tracking que permita detectar las parcelas del cultivo y a partir de ello poder determinar cuándo debe sensar el cultivo. Con la particularidad que el tracking será invertido, ya que el objeto a seguir (planta) estará quieto y lo que se mueve es la cámara.

Por último, la plataforma trabajará de modo conjunto con un vehículo aéreo no tripulado, al cual se necesita proveer de sistema de visión que permita realizar un mapeo del cultivo, pudiendo ésta detectar parcelas, tanto para navegar el cultivo como para además, a partir de las imágenes obtenidas, obtener información sobre el cultivo.

Resultados y Objetivos

Se espera dotar de sistemas de visión

artificial a una plataforma robótica móvil con un sistema de navegación autónoma en escenas outdoor, y en condiciones de suelo irregular, que pueda aplicarse al recorrido de ensayos de diferentes tipos de cultivo en el campo. Sumado a un sistema de visión para un vehículo aéreo no tripulado. Donde ambos sistemas puedan estar relacionados entre sí y que sirvan de aporte el uno del otro. Trabajando de manera conjunta en el sensado de datos del cultivo.

Debido a que la UNNOBA se encuentran ubicada dentro de la Región Pampeana, más específicamente en lo que se denomina la Pampa Húmeda, una de las regiones más relevantes en lo que respecta a producción agrícola, se espera que este trabajo sea el puntapié de la creación de una plataforma que sirva de soporte a los especialistas del sector agrícola. Sumado al beneficio directo de los recursos humanos involucrados, con el objetivo fundamental de consolidar un equipo de trabajo que se especialice en la temática.

Formación de Recursos Humanos

En esta línea de I/D se espera concluir con un trabajo de Tesina de Grado de la carrera de Licenciatura en Sistemas, sumado a la formación de tres alumnos becarios de la universidad que se encuentran cursando la carrera de Ingeniería en Informática. Además la realización de dos Tesis de posgrado, y una Beca de Estudio Cofinanciada otorgada por la **Comisión de**

Investigaciones Científicas (CIC) y la UNNOBA.

Bibliografía

- [1] E. R. and J. L. Y. Wang, K. Lee, S. Cui, “Agriculture robot and applications,” in Future information engineering and manufacturing science : proceedings of the 2014 International Conference on Future Information Engineering and Manufacturing Science (FIEMS 2014), 2015, pp. 43–46.
- [2] “Sydney University - Australian Centre for Field Robotics.” [Online]. Available: <https://sydney.edu.au/engineering/our-research/robotics-and-intelligent-systems/australian-centre-for-field-robotics/agriculture-and-the-environment.html>. [Accessed: 09-Mar-2018].
- [3] T. Mueller-Sim, M. Jenkins, J. Abel, and G. Kantor, “The Robotanist: A ground-based agricultural robot for high-throughput crop phenotyping,” in 2017 IEEE International Conference on Robotics and Automation (ICRA), 2017, pp. 3634–3639.
- [4] C. Paper, P. Biber, and R. Bosch, “Field-Robot-Based Agriculture: ‘RemoteFarming.1’ and ‘BoniRob-Apps,’” in AgEng

- 2013, 71. Internationale Tagung Land.Technik, 2013, no. November, pp. 439–445.
- [5] “Agco/Fendt - MARS.” [Online]. Available: <https://www.fendt.com/es/fendt-mars.html>. [Accessed: 09-Mar-2018].
- [6] “Agco/Fendt - MARS.” [Online]. Available: <https://www.fendt.com/es/fendt-mars.html>. [Accessed: 09-Mar-2018]
- [7] “DJI - Farm Drones.” [Online]. Available: <https://enterprise.dji.com/agriculture?site=brandsite&from=nav>. [Accessed: 09-Mar-2018].
- [8] S. Rossius, “Reconocimiento de objetos mediante webcam en tiempo real”. Proyecto Final de Carrera. Escuela Politécnica Superior de Gandía Ingeniería Técnica de Telecomunicaciones. Universidad Politécnica de Valencia. 2009.
- [9] P. Schiaffino, “Navegación autónoma basada en visión estereoscópica y lógica difusa”. Universidad Nacional de Luján. 2006
- [10] D.G. Gomez, “Visión artificial aplicada a vehículos inteligentes”. Universidad Autónoma de Barcelona. 2004
- [11] S. L. Sandoval Niño and F. A. Prieto Ortiz, “Caracterización de granos de café cereza empleando técnicas de visión artificial”. Facultad de Ingeniería, Universidad Pedagógica y Tecnológica de Colombia. 2007.
- [12] [1] J. M. Guerrero Hernández, “Sistema De Visión Para Agricultura De Precisión: Identificación En Tiempo Real De Líneas De Cultivo Y Malas Hierbas En Campos De Maíz,” Universidad Complutense de Madrid, 2015.
- [13] J. L. Araus and J. E. Cairns, “Field high-throughput phenotyping: The new crop breeding frontier,” *Trends Plant Sci.*, vol. 19, no. 1, pp. 52–61, 2014.
- [14] D. R. Laorden, “Obtención de distancias y posiciones geográficas mediante técnicas de visión artificial”. Facultad de Informática. Universidad Complutense de Madrid. 2008
- [15] J. C. Mejía Ospina - F. A. Prieto Ortiz - J. W. Branch Bedoya, “Corrección de iluminación para imágenes aéreas de cultivos tomadas a baja altitud”. Universidad Nacional de Colombia. 2007.

Sensores: Pilares de la investigación tecnológica

Russo C.^{1,2}, Sarobe M.¹, Alonso N.¹, Alvarez E.¹, Charne J.¹, Di Cicco A.¹, Jaszczyszyn A.¹, Jatip N.³, Llanos E.³, Luengo P.¹, Moran M.³, Moretti N.³, Muscia A.³, Osella Massa G., Ramon H.^{1,2}, Schiavoni A.³, Terzano J.³, Beloso J.³, Natal M.³

Instituto de Investigación y Transferencia de Tecnología (ITT)⁴
 Comisión de Investigaciones Científicas (CIC)
 Escuela de Tecnología (ET)
 Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

Sarmiento y Newbery, Junín (B) – TE: (0236) 4477050

{claudia.russo, monica.sarobe, eduardo.alvarez, javier.charne, carlos.dicicco, adrian.jaszczyszyn, nicolas.jatip, emanuel.llanos, pablo.luengo, matias.moran, nicolas.moretti, german.osella, hugo.ramon, antonela.schiavoni, julieta.terzano, .juanpablo.beloso, matias.natal}@itt.unnoba.edu.ar

Resumen

En el transcurso de diferentes proyectos realizados en el ITT / UNNOBA, nos hemos dado cuenta que resulta difícil encontrar el dispositivo transductor o sensor adecuado desde un inicio, haciendo necesario producir más de un prototipo de un mismo trabajo.

Los proyectos descritos en este documento tienen como propósito identificar, contextualizar, evaluar y permitir la toma de decisiones vinculada a la temática abordada. Mencionaremos los distintos sensores en los que estamos trabajando como: cardíaco, acelerómetro, humedad, temperatura, ultrasonido, infrarrojo y captura de imágenes entre otros.

Someter las ideas a proyectos concretos y tener que lidiar con factores como la luz natural, temperatura, humedad,

desplazamiento de cuerpos y límites concretos, nos permitió experimentar avances y frustraciones que debieron ser resueltas con replanteos y cambios de rumbo en las investigaciones. En muchos casos el sensor de mayor costo y precisión no es necesariamente el más beneficioso para lograr los resultados deseados.

En este contexto, la selección de los sensores apropiados en los proyectos de investigación con base tecnológica, son claves para el éxito de los mismos, estrategia que trata esta línea de investigación.

Palabras clave: sensores, prestaciones, datos, precisión, decisiones.

¹ Docente Investigador en el Instituto de Investigación y Transferencia en Tecnología (ITT) / Escuela de Tecnología / UNNOBA

² Investigador Asociado Adjunto sin director a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

³ Becario PROMING / ITT / Escuela de Tecnología

⁴ Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

Contexto

En el marco del proyecto de investigación "Informática y tecnologías emergentes" que se desarrolla en la UNNOBA, con lugar de trabajo en el Instituto de Tecnología y Transferencia (ITT) se comienza a trabajar con esta línea de investigación orientada al trabajo con sensores aplicado a campos como el deporte, agroindustria, arte digital, educativo entre otros.

1- Introducción

Hoy en día nos encontramos con variados equipos y dispositivos que cuentan con aplicaciones que ya incorporan diferentes tipos de sensores que miden varias magnitudes físicas, dando cuenta por ejemplo de los distintos movimientos realizados, su orientación, inclinación, hasta medio ambientales.

Los datos que proporcionan algunas magnitudes físicas como el giro en tres ejes, midiendo las fuerzas de aceleración como acelerómetros, giróscopos, de rotación, GPS nos dará cuenta de los movimientos realizados por el dispositivo o que/quien lleve esos sensores.

Así también las magnitudes medio ambientales de temperatura, humedad, presión, nivel de iluminación, contaminación acústica serán aquellos elementos como termómetros, barómetros, fotómetros que nos darán datos que nos serán útiles a posterior para obtener información ambiental.

Ahora bien es un desafío encontrar y obtener las magnitudes físicas que proporcionan los sensores, cuentan con la rango, precisión, exactitud, sensibilidad, consumo de energía, etc; necesarias para que las aplicaciones sean útiles para los distintos proyectos abocados. Para lo cual se considera de importancia definir estos parámetros.

Transductores

Es un dispositivo capaz de transformar o convertir un determinado tipo de energía de entrada, en otra de diferente a la salida. El transductor sirve para obtener la información de entornos físicos y químicos y conseguir señales o impulsos eléctricos o viceversa. Los transductores siempre consumen algo de energía por lo que la señal medida resulta debilitada.

Como parte del transductor está incluido el sensor y el actuador o electrónica necesaria para entregar la señal en un formato electrónico deseado, como voltaje o cualquier estándar industrial.

Sensores

Se entiende por sensores como un dispositivo apto para detectar acciones o estímulos externos y responder en consecuencia. Estos aparatos pueden transformar las magnitudes físicas o químicas en magnitudes eléctricas. [1] [9][11][12]

Características técnicas de un sensor:

Descripción en régimen permanente:

- **Rango:** valores máximos y mínimos para las variables de entrada y salida de un sensor.
- **Exactitud:** la desviación de la lectura de un sistema de medida respecto a una entrada conocida. El mayor error esperado entre las señales medida e ideal.
- **Repetitividad:** la capacidad de reproducir una lectura con una precisión dada.
- **Reproducibilidad:** tiene el mismo sentido que la repetitividad excepto que se utiliza cuando se toman medidas distintas bajo condiciones diferentes.
- **Resolución:** la cantidad de medida más pequeña que se pueda detectar.

- **Error**: es la diferencia entre el valor medido y el valor real.

- **No linealidades**: la desviación de la medida de su valor real, supuesto que la respuesta del sensor es lineal. No-linealidades típicas: saturación, zona muerta e histéresis.

- **Sensibilidad**: es la razón de cambio de la salida frente a cambios en la entrada: $s = \partial V / \partial x$

- **Excitación**: es la cantidad de corriente o voltaje requerida para el funcionamiento del sensor.

- **Estabilidad**: es una medida de la posibilidad de un sensor de mostrar la misma salida en un rango en que la entrada permanece constante.

- **Ruido**.

Descripción en régimen dinámico:

- **Tiempo de retardo**: t_d , es el tiempo que tarda la salida del sensor en alcanzar el 50% de su valor final.

- **Tiempo de subida**: t_r , es el tiempo que tarda la salida del sensor hasta alcanzar su valor final. => velocidad del sensor, es decir, lo rápido que responde ante una entrada.

- **Tiempo de pico**: t_p , es el tiempo que tarda la salida el sensor en alcanzar el pico máximo de su sobreoscilación

- **Pico de sobreoscilación**: M_p , expresa cuánto se eleva la evolución temporal de la salida del sensor respecto de su valor final.

- **Tiempo de establecimiento**: t_s , el tiempo que tarda la salida del sensor en entrar en la banda del 5% alrededor del valor final y ya no vuelve a salir de ella. [2], [3]

Sensores utilizados en proyectos de Investigación en el ITT / UNNOBA:

A continuación en la Fig.1 haremos mención y síntesis de los sensores y sus controladores

utilizados en cuatro proyectos /casos diferentes.

En los primeros tres, se está utilizando distintos tipos de sensores conectados a un microcontrolador Atmega 328 [9] y en el último se trabaja con la placa EDU-CIAA [4] con su microcontrolador LPC4337 (dual core ARM Cortex-M4F y Cortex-M0).

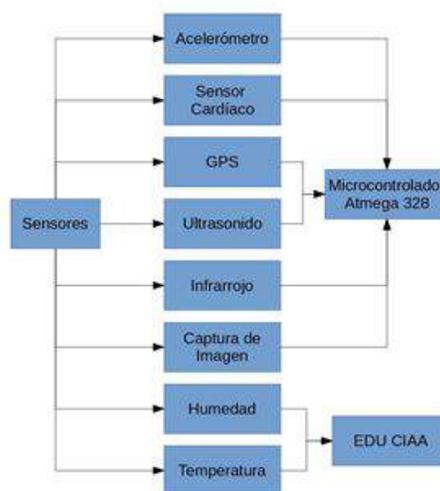


Fig.1: Sensores y controladores

Proyecto “GPS aplicado al deporte”. El primer prototipo está formado por 5 módulos diferentes: microcontrolador, módulo GPS, sensor cardíaco, acelerómetro [10] y módulo bluetooth. Los elementos que permiten una entrada de datos al microcontrolador para el posterior procesamiento son: módulo GPS, sensor cardíaco y acelerómetro.

El GPS es el encargado de indicar en intervalos de tiempo muy pequeños la latitud y longitud que determina la ubicación del jugador. Al utilizar módulos de posicionamiento que no están orientados para uso militar, se deben establecer estrategias

para minimizar el error de ubicación. Con estos datos sumados a los del acelerómetro y el sensor cardíaco. Para la extracción y análisis de los datos recolectados por el dispositivo se desarrollará una aplicación web dinámica, que permita tomar decisiones en base a las distintas formas de visualización y análisis de los datos.

Proyecto “REP – Robot Educativo Programable”, es el desarrollo con más horas de trabajo y donde se tuvieron que realizar 4 versiones del robot. En la Versión 2.0 se usaban sensores infrarrojos en el frente del robot para el avance motriz pero al realizar pruebas fuera del laboratorio comprobamos que el comportamiento no era confiable por la incidencia de la luz natural. Se solucionó el problema cambiando sensores infrarrojos por los de ultrasonido.

Proyecto “Scanner 3D”. Con el objetivo de sumar nuevo equipamiento económico para el Laboratorio CNC, se encontró un Scanner 3d de código libre llamado “Ciclop” donde el sensor fundamental es una web-cam combinado con 2 laser’s de baja potencia para reconstruir la figura colocada en el plato giratorio. El desarrollo del proyecto no lo realizamos nosotros por ende no seleccionamos ni el sensor, ni los actuadores pero tenemos planeado hacer variaciones del modelo después de lograr su funcionamiento.

Proyecto “Visualización de variables ambientales con Edu-CIAA”. Para el campo experimental “Las Magnolias” de la UNNOBA es necesario sensor y tomar registro de las variables de temperatura y humedad del ambiente. Por tal motivo, se utilizó la placa EDU-CIAA [4] que posee características de múltiples entradas y salidas tanto analógicas como digitales, las que pueden permitir el estudio de conectividad

con distintos tipos de sensores para el registro de las variables a sensor de acuerdo a los requerimientos necesarios.

Los sensores están conectadas a dos entradas del tipo analógica que contienen conversores analógicos digitales. Utilizando estos periféricos que contiene el microcontrolador es como se digitalizan las señales de temperatura y humedad provistas por los sensores. El lenguaje de programación es micropython [7] que permite en pocas líneas de código ejecutar operaciones complejas, como lo son leer los valores obtenidos por los sensores y mostrarlos en un display LCD. En un segundo prototipo se utiliza otro tipo de sensor, el DTH22 [8] que también es llamado AM2302 [8] que sensa temperatura y humedad y lo entrega al microcontrolador de manera ya digital sin tener que agregar componentes adicionales para acondicionar la señal para lograr una medición.

2- Líneas de Investigación y Desarrollo

La línea de investigación de estos proyectos es el uso de las nuevas herramientas y prototipos para poder evaluar los distintos comportamientos de los sensores y determinar los posibles condicionamientos o limitaciones que tienen para obtener los datos de manera confiable y que sean de calidad para futuras interpretaciones, haciendo hincapié en los siguientes ejes;

1. Análisis del entorno para definir y contextualizar la temática que debemos abordar.
2. Estudio de los dispositivos que necesitamos en base a la disponibilidad del mercado.

3. Definición de estrategias de producción de prototipos, para poder realizar los test y mediciones correspondientes.

4. Selección del modelo más eficiente para el problema planteado.

3- Resultados Obtenidos/Esperados

Esperamos que la aplicación de nuestros proyectos en distintas situaciones nos permita generar una matriz de conocimiento donde podamos ver en cada uno de los escenarios propuestos el tipo de sensores utilizados y en base a nuestra investigación y pruebas de campo cual consideramos más adecuado.

Deseamos difundir y transferir los logros alcanzados mediante la presentación y participación en diferentes congresos, jornadas y workshops de carácter nacional e internacional relacionados con la temática.

4- Formación de Recursos Humanos

El equipo de trabajo está compuesto por investigadores del área de Sistemas y del área de Ciencias de la Educación, se cuenta con investigadores en formación, becarios de posgrados y becarios de grado. Además se incorporan alumnos para que se inicien en tareas de investigación.

Se espera contribuir para generar 2 Tesinas de Licenciatura en Sistemas, 1 Prácticas Profesionales Supervisadas llamada "Visualización de variables ambientales con la EDU-CIAA", 1 Becario de intercambio internacional.

Actualmente se encuentran en desarrollo 1 trabajo de Especialista, 4 Tesis de Magíster y 1 Tesis Doctoral.

5- Bibliografía

- [1] Pallás Areny, R, (2003), *Sensores y Acondicionadores de señal*, Barcelona- España, Boixareu.
- [2] http://www.logicbus.com.mx/transductores_sensores.php
- [3] https://www.edu.xunta.gal/centros/iescamposanalberto/aulavirtual2/pluginfile.php/9022/mod_resource/content/0/Sistemas_de_Control.pdf
- [4] Edu-CIAA. [OnLine]: <http://www.proyecto-ciaa.com.ar/dev/wiki/doku.php?id=desarrollo:edu-ciaa:edu-ciaa-nxp>
- [5] Sensor de temperatura lm35 [OnLine]: <http://www.ti.com/lit/ds/symlink/lm35.pdf>
- [6] Sensor de humedad relativa HIH-4000-00 [OnLine]: <http://sensing.honeywell.com/HIH-4000-001-Humidity-Sensors>
- [7] Portal de micropython [OnLine]: <https://micropython.org/>
- [8] Sensor DTH22 . AM2302 [OnLine]: <https://www.sparkfun.com/datasheets/Sensors/Temperature/DHT22.pdf>
- [9] Ramírez, L., Jiménez, G., Carreño, J. (2014), *Sensores y actuadores: Aplicaciones con Arduino*, Grupo Editorial Patria
- [10] Kymberly Tuck., "Tilt sensing using linear accelerometers", *Acelerometer Systems and Applications Systems. Rev. 6*, 2007.
- [11] Fernández, C., García, A.(2009), *Sensores y actuadores*, Departamento de Publicaciones de la Escuela Universitaria de Informática de Madrid
- [12] Alvarez, H., Robledo Fernández, D., Bretón, A. (2014), *Desarrollo de aplicaciones para Android II*, Ministerio de Educación, Cultura y Deporte. Área de Educación

Seguridad Infomática

Voto Electrónico Seguro con Criptografía Homomórfica

Pablo García ¹; Jeroen van de Graaf ²; Germán Montejano ³;

¹ Departamento de Matemática
Universidad Nacional de La Pampa
Av. Uruguay 151 – (6300) Santa Rosa – La Pampa – Argentina
Tel.: +54-2954-245220– Int. 7125
pablogarcia@exactas.unlpam.edu.ar – web: <http://www.exactas.unlpam.edu.ar>

² Departamento de Ciência da Computação
Universidade Federal de Minas Gerais
Av. Antonio Carlos, 6627 – 31270-010 - Belo Horizonte – Minas Gerais - Brasil
Tel.:+55-3409-5836
jvdg@dcc.ufmg.br – web: <http://www.dcc.ufmg.br/~jvdg>

³ Departamento de Informática
Universidad Nacional de San Luis
Ejército de los Andes 950 – (5700) San Luis – San Luis – Argentina
Tel.: +54-266-4520300– Int. 2128
gmonte@unsl.edu.ar – web: <http://www.unsl.edu.ar>

Resumen

La viabilidad de la implementación del voto electrónico es un tema extremadamente discutido en la República Argentina. Por ejemplo, el Congreso de la Nación ha rechazado la propuesta del Poder Ejecutivo Nacional de implementar un sistema de boleta única electrónica, en concordancia con una publicación de un grupo de informáticos pertenecientes a universidades nacionales argentinas¹.

Se considera, sin embargo, que no existen argumentos contundentes para afirmar que es imposible conseguir que un sistema de votación electrónica garantice transparencia y privacidad para el votante. De hecho, existen sólidas propuestas de tipo híbrido que propor-

cionan evidencia física que permite asegurar que los procedimientos se llevaron a cabo de manera irreprochable². Ese punto es crucial: el sistema debe demostrar de forma clara e indiscutible la transparencia del proceso de manera tal que cualquier ciudadano común pueda verificar los resultados y a la vez, convencerse de que el proceso se llevó a cabo de manera totalmente confiable.

En consecuencia, se propone analizar el problema en profundidad y generar un modelo que proporcione respuestas apropiadas. Además de la descripción detallada de la criptografía propuesta se llevará a cabo la implementación de un prototipo que implemente las funcionalidades básicas.

Palabras clave: *Voto Electrónico, Transparencia, Anonimato, Evidencia*

¹ <http://www.cronista.com/economiapolitica/Expertos-universitarios-lanzaron-una-campa-na-contr-a-el-voto-electronico-20161101-0113.html>

² <https://www.usenix.org/conference/evt-wo-te13/workshop-program/presentation/bell>

Física, Verificabilidad E2E, Criptografía Homomórfica, Paillier, ElGamal.

Contexto

Por Resolución N° 488/14 del Consejo Directivo de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa se acredita el Proyecto de Investigación: "Aspectos de Seguridad en Proyectos de Software", que se desarrolla en el ámbito de la FCEyN de la UNLPam. El mismo es dirigido por el Doctor Germán Antonio Montejano y codirigido por el Magister Pablo Marcelo García e incluye a los magisters Silvia Gabriela Bast y Daniel Vidoret como investigadores.

El Proyecto surge desde la línea de Investigación "Ingeniería de Software y Defensa Cibernética", presentada en [1], y que a su vez se enmarca en el Proyecto "Ingeniería de Software: Aspectos de Alta Sensibilidad en el Ejercicio de la Profesión de Ingeniero de Software" de la Facultad de Ciencias Físico - Matemáticas y Naturales de la Universidad Nacional de San Luis (UNSL, <http://www.sel.unsl.edu.ar/pro/proyec/2012/index.html>) y que incluye acciones de cooperación con la Universidad Federal de Minas Gerais (UFMG, Brasil).

Entre tales acciones debe mencionarse que Jeroen van de Graaf, PhD., Docente de UFMG, y el Dr. Germán Montejano (UNSL) fueron orientadores del Mg. Pablo García en el desarrollo de su tesis de Maestría en Ingeniería de Software titulada "Optimización de un Protocolo Dining Cryptographers Asíncrono", defendida en 2013 en el ámbito de la UNSL. Durante el desarrollo de la misma se generaron una serie de publicaciones de avances parciales, como por ejemplo [2], [3], [4], [5] y [6].

Introducción

La implementación del voto electrónico es un tema muy discutido en la actualidad. Existen partidarios y detractores, en proporciones similares. Lo que debe quedar claro es que la implementación de un modelo innovador debe, necesariamente, resultar superior de las propuestas preexistentes. De no ser así no tiene sentido ninguna propuesta novedosa.

En consecuencia, la implementación de sistemas de voto electrónico presenta dos exigencias claras:

- El escrutinio asociado debe reflejar de manera indiscutible la voluntad de los ciudadanos
- Simultáneamente, los electores deben estar convencidos de que su privacidad es garantizada indefinidamente.

Se otorga máxima atención a las consecuencias que puede acarrear, para cualquier votante, el hecho de que su opción sea divulgada. Existen numerosas variantes de prácticas deshonestas que se derivan de conocimiento de esa información, con o sin el aval del elector. Es claro que si un ciudadano pudiera probar que votó a un determinado partido político, podría recibir una contraprestación. Análogamente, si un sector detecta que un votante votó a otra opción, podría llevar a cabo acciones que perjudiquen al votante.

Cualquier propuesta que se desee implementar en ese sentido, deberá cumplir con una serie de condiciones que se exigen actualmente a los sistemas de votación electrónica, [7]. Las principales son las siguientes:

- Debe existir evidencia física que garantice la transparencia del proceso [8]. Este punto debe ser implementado

de tal manera que sirva como prueba irrefutable, no sólo en la consideración de los expertos, sino también de los electores.

- Utilización de métodos criptográficos cuya seguridad pueda ser demostrada de manera matemática y que garanticen el anonimato del votante y la seguridad computacional necesaria en lo referido a la transparencia de los resultados de los comicios. En ese sentido, revisten máximo interés los esquemas homomórficos, es decir aquellos que permiten operaciones matemáticas sobre las versiones cifradas de la información. Ejemplos de modelos que presentan esa característica son Paillier [9], [10] y ElGamal exponencial [11], [12].

- Aplicación del concepto de independencia del software, es decir que si el software es corrupto, no hay ninguna posibilidad de que se generen resultados incorrectos y eso no sea detectado [13].

- Definir un modelo concreto para la aplicación de verificabilidad “End to End” (E2E) [14], [15]. Ello implica garantizar:

- Verificabilidad individual: es decir, que cualquier votante puede asegurarse de que su voto fue correctamente contabilizado.

- Verificabilidad universal, que implica que cualquier observador neutral pueda asegurarse de que todos los sufragios han sido correctamente contabilizados.

- Verificabilidad de padrón, de manera tal que sea posible probar que todos los votos incluidos en el recuento provienen de votantes habilitados.

- Imposibilidad de que algún votante pueda demostrar por quién votó, dado que eso daría lugar a maniobras de “compra de sufragios”.

Líneas de Investigación, Desarrollo e Innovación

El grupo de trabajo investiga, básicamente sobre dos líneas paralelas, para generar modelos que pudieran aplicarse a los sistemas de voto electrónico:

- Basados en criptografía One Time Pad, desarrollada por la Magister Silvia Gabriela Bast.
- Basados en criptografía homomórfica, que es la línea relacionada con este documento y sus autores.

Resultados y Objetivos

En el ámbito del voto electrónico, este grupo de trabajo ha realizado durante el año 2017, las siguientes publicaciones:

- [16] expone una técnica de recuento y recuperación de sufragios para el modelo OTP – Vote cuyos conceptos pueden generalizarse a otros esquemas.
- [17] especifica una propuesta para el modelo de datos aplicable al sistema OTP – Vote.
- [18] propone un modelo para agregar verificabilidad OTP – Vote.
- [19] presenta una técnica para generar una optimización en el almacenamiento de sufragios en esquemas basados en el protocolo Non Interactive Dining Cryptographers (NIDC).
- [20] Se expone un sistema de generación de códigos para el sistema OTP – Vote, que optimiza el dicho proceso de manera significativa.

Llegado este punto, se decide investigar un camino alternativo, relacionado con la criptografía homomórfica. En

consecuencia, a futuro, se pretende llevar a cabo las siguientes acciones:

- Proponer un nuevo esquema de voto electrónico basado en las características expuestas en las secciones previas.
- Implementar una aplicación experimental que permita observar el comportamiento del modelo teórico que surja del avance de la investigación.
- Continuación del relevamiento de aplicaciones orientadas al voto electrónico, con el fin de detectar fallencias y proponer mejoras.

Formación de Recursos Humanos

En el marco del presente proyecto se presentan los siguientes puntos relacionados con la formación de recursos humanos durante 2017:

- Pablo García realizó una estadía de un mes (3/9/2017) al 4/10/2017) en la Universidade Federal de Minas Gerais (UFMG), en el Departamento de Ciência da Computação (DCC) dependiente del Instituto de Ciências Exatas (ICEX) trabajando en el laboratorio 4303 del grupo “Criptografía Teórica y Aplicada”, dirigido por Jeroen van de Graaf, PhD.
- Pablo García completó el cursado de la totalidad de los créditos exigidos en el Doctorado en Ingeniería Informática en la Facultad de Ciencias Físico Matemáticas y Naturales de la Universidad Nacional de San Luis (UNSL).
- Pablo García presentó su Plan de Tesis Doctoral, en el marco del

Doctorado en Ingeniería Informática en la Facultad de Ciencias Físico Matemáticas y Naturales de la Universidad Nacional de San Luis (UNSL). El mismo se encuentra en proceso de evaluación. Dicha tesis será dirigida por Jeroen van de Graaf, PhD. y el Dr. Germán Montejano.

Referencias

[1] Uzal R., van de Graaf J., Montejano G., Riesco D., García P.: “Inicio de la Línea de Investigación “Ingeniería de Software y Defensa Cibernética”. Memorias del XV Workshop de Investigadores en Ciencias de la Computación 2013 (WICC 2013). Ps.769 - 773. ISBN: 9789872817961. 2013.

[2] van de Graaf J., Montejano G., García P.: “Optimización de un Protocolo Non-Interactive Dining Cryptographers”. Congreso Nacional de Ingeniería Informática / Sistemas de Información. CoNaIISI 2013. Córdoba, Argentina.

[3] van de Graaf J., Montejano G., García P., Bast S.: “Anonimato en Sistemas de Voto Electrónico”. Memorias del XVI Workshop de Investigadores en Ciencias de la Computación 2014 (WICC 2014). Ps. 822 – 826. ISBN: 9789503410844. 8 y 9 de mayo de 2014.

[4] van de Graaf J., Montejano G., García P.: “Manejo de Colisiones en un Protocolo Non Interactive Dining Cryptographers”. Anales de las 42° Jornadas Argentinas de Informática e Investigación Operativa (JAI-IO, ISSN: 1850-2776). Workshop de Seguridad Informática (WSegI 2013, ISSN: 2313-9110). Ps. 29 a 43. Septiembre 2013.

[5] García P., van de Graaf J., Montejano G., Bast S., Testa O.: “Implementación de Canales Paralelos en un Protocolo Non Interactive Dining Cryptographers”. 43° Jornadas Argentinas de Informática e Inves-

- tigación Operativa (JAIIO 2014), Workshop de Seguridad Informática (WSegI 2014).
- [6] García P., van de Graaf J., Hevia A., Viola A.: “Beating the Birthday Paradox in Dining Cryptographer Networks”. The third International Conference on Cryptology and Information Security in Latin America, Latincrypt 2014. September 17-19, 2014. Florianopolis, Brasil. Lecture Notes in Computer Science, Springer (2014).
- [7] Hao, F, Ryan P.: “Real -World Electronic Voting. Design, Analysis And Deployment”. Cr Press. ISBN-13: 978-1498714693. ISBN-10: 1498714692. 2017.
- [8] Prince, A.: “Consideraciones, Aportes y Experiencias para el Voto Electrónico en Argentina”. Editorial Dunken. ISBN: 978-987-02-1732-9. 2006.
- [9] Volkhausen T.: “Paillier Cryptosystem: A Mathematical Introduction”. 2006.
- [10] O’Keeffe M.: “The Paillier Cryptosystem: A Look Into The Cryptosystem And Its Potential Application”. The College of New Jersey Mathematics Department. 2008
- [11] El Gamal T. “A public key cryptosystem and a signature scheme based on discrete logarithms”. In Proceedings of CRYPTO 84 on Advances in cryptology, pages 10–18. Springer-Verlag New York, Inc. 1985.
- [12] Koscielny C.: “A New Approach to the Elgamal Encryption Scheme”. Academy of Management of Legnica, Faculty of Computer Science. 2004.
- [13] Rivest R.: “On the notion of ‘software independence’ in voting systems”. Philosophical Transactions of The Royal Society A, 366(1881):3759–3767. 2008.
- [14] Benaloh J. Bernhard M. Halderman J. Rivest R Ryan P. Stark P. Vora P. Teague V. Wallach D.: “Public Evidence from Secret Ballots”. Documento presentado en E-Vote-ID 2017.
- [15] Kelsey J., Regenscheid A., Moran T., Chaum D.: “Attacking Paper-Based E2E Voting Systems”. In: Chaum D. et al. (eds) Towards Trustworthy Elections. Lecture Notes in Computer Science, vol 6000. Springer, Berlin, Heidelberg. ISBN: 978-3-642-12979-7. 2010.
- [16] García P., Bast S., Montejano G.: “Recuento y Recuperación de Sufragios en OTP – Vote”. Simposio de Informática en el Estado (SIE) del XLIII CLEI (Conferencia Latinoamericana de Informática) y 46° JAIIO (Jornadas Argentinas de Informática e Investigación Operativa. ISSN: 2451-7534. Ps. 38 a 51.
- [17] Bast S., García P., Montejano G.: “Modelo de Datos del Sistema de Voto Electrónico Presencial OTP-Vote”. Simposio de Informática en el Estado (SIE) del XLIII CLEI (Conferencia Latinoamericana de Informática) y 46° JAIIO (Jornadas Argentinas de Informática e Investigación Operativa). ISSN: 2451-7534. Ps. 23 a 37. 2017.
- [18] García P., Bast S., Montejano G.: “Verificabilidad ‘End to End’ para OTP – Vote”. VI Workshop Seguridad Informática (WSI) del Congreso Argentino de Ciencias de la Computación (CACIC 2017).
- [19] García P., Bast S., Montejano G.: “Efficient Votes Storage in a Non-Interactive Dining Cryptographers (NIDC) Protocol”. VI Workshop Seguridad Informática (WSI) del Congreso Argentino de Ciencias de la Computación (CACIC 2017).
- [20] Bast S., García P., Montejano G.: “Generación de Códigos para OTP – Vote”. 5° Congreso Nacional de Ingeniería Informática / Sistemas de Información. CONAIISI 2017. [http://tecnomate .xyz/Actas-CONAIISI-2017.pdf](http://tecnomate.xyz/Actas-CONAIISI-2017.pdf). Ps. 12 a 22.

Criptografía Liviana para aplicar en dispositivos IoT

Mg. Jorge Eterovic; Esp. Marcelo Cipriano;

Instituto de Investigación en Ciencia y Tecnología
Dirección de Investigación Vicerrectorado de Investigación y Desarrollo.
Universidad del Salvador.
Lavalle 1854 – C1051AAB -Ciudad Autónoma de Buenos Aires - Argentina

jorge.eterovic@gmail.com – marcelo.cipriano@usal.edu.ar

RESUMEN

En los últimos tiempos se ha observado un notorio aumento en la cantidad y calidad de la conectividad de los dispositivos usados por las personas y hasta en aquellos que conectan dispositivos entre sí. Entre las múltiples razones de ello se encuentra lo que se ha dado en llamar *MtoM*¹ (o también *M2M*). Una de sus áreas es conocida por el nombre de “*Internet de las Cosas*”², la cual permite la conexión de objetos de distinta naturaleza e índole, a través de Internet. También las redes conocidas como *WSN*³ y dispositivos *RFID*⁴ se suman al ecosistema. La incidencia de estos factores sobre la humanidad se vislumbra como un cambio de paradigma.

Esta “nueva era” de la humanidad, en la que enormes cantidades de información son transmitidas y procesadas, conlleva un enorme desafío: la seguridad de la misma. Dados los perfiles de hardware y software de muchos de estos dispositivos, existe la posibilidad que no se cuenten con los mecanismos de seguridad adecuados.

Desde datos médicos en tiempo real obtenidos por dispositivos e-Health⁵, chips subcutáneos para la identificación y rastreo de personas y

zapatillas GPS[1] los dispositivos transmiten y reciben información que dejan expuesto al usuario a riesgos no sólo contra la confidencialidad de sus datos sino también al alcance de otros delitos.

Es por ello que este proyecto propone la realización de un estudio y análisis de algoritmos criptográficos -que podrían ser ejecutados en dispositivos con limitados recursos de hardware y software- haciendo uso de Criptografía Ligera[2].

Palabras Clave:

Criptografía Ligera, RFID, Internet de las Cosas, Internet of Things.

CONTEXTO

El Vicerrectorado de Investigación y Desarrollo (VRID), perteneciente a la Universidad Nacional del Salvador (USAL), dicta las políticas referidas a la investigación, concibiéndola como un servicio a la comunidad, entendiendo que los nuevos conocimientos son la base de los cambios sociales y productivos. Con el impulso de las propias Unidades Académicas se han venido desarrollando acciones conducentes a concretar proyectos de investigación uni/multidisciplinarios, asociándolos a la docencia de grado y postgrado y vinculando este accionar, para potenciarlo, con otras instituciones académicas del ámbito nacional e internacional.

La Dirección de Investigación, dependiente del VRID, brinda soporte a las distintas Unidades de Investigación de la y a sus investigadores para el desarrollo de Proyectos y Programas de Investigación, nacionales e

¹ Machine to Machine: máquina a máquina. Se refiere la comunicación para el intercambio de información entre dos dispositivos distantes o remotos.

² Internet of Things: Internet de las Cosas.

³ Wireless Sensor Network: Redes Inalámbricas de Sensores.

⁴ Radio Frequency Identification: identificación por radiofrecuencia.

⁵ E-Salud: cuidados sanitarios apoyados en dispositivos TIC's como pueden ser marcapasos, bombas de insulina, implantes cocleares, etc.

internacionales, como así también, apoyo y orientación de recursos para la investigación. A ella pertenece el Instituto de Investigación en Ciencia y Tecnología (RR 576/12) en el cual se enmarca este proyecto, con una duración de 2 años (2017-2018).

1. INTRODUCCIÓN

La llamada *Internet de las Cosas* promete un cambio sin igual en la historia humana[3] que afectará directa o indirectamente a campos tan importantes, como son:

- Cuidados médicos.
- Manufactura de productos.
- Uso de la energía.
- Infraestructura urbana.
- Seguridad.
- Extracción de recursos naturales.
- Agricultura.
- Ventas.
- Vehículos.

A su vez, estos cambios se apoyan en tecnologías y dispositivos que están limitados en recursos, dada su naturaleza, como son entre otros:

- Espacio.
- Consumo de energía.
- Almacenamiento en Memoria.
- Capacidad de cómputo.

Existe además una tendencia a aumentar la cantidad de dispositivos que requieran conexiones a Internet. En particular datos relevados en el *Ericsson Mobility Report* del año 2015 prevee que 28.000.000.000⁶ de teléfonos estarán conectados para el año 2021, más de la mitad de ellos con capacidades de IoT y M2M[4].

Dicha empresa, en el último reporte del año 2017 se informa que el tráfico de datos creció 65% interanual entre 2016 y 2017[5] y que 1.800.000.000 de dispositivos IoT obtendrán conectividad a través de teléfonos celulares para el año 2023.

⁶ Asumiendo una población mundial de alrededor de 7.000.000.000 personas, este valor indica una media de 4 teléfonos por persona en el planeta.

Esta demanda de conectividad será satisfecha con la nueva tecnología 5G⁷. Estos dispositivos móviles intercambiarán información con objetos de la vida cotidiana: desde zapatillas con GPS, marcapasos, heladeras que elaboran listas de compras y demás[6]. La lista no deja de crecer año a año. Estos dispositivos y otros que se sumarán a los existentes, comparten la imperiosa necesidad de asegurar la información que procesan y transmiten. Pero por su propia naturaleza, tienen limitaciones de Hardware y Software que impiden el uso de mecanismos criptográficos tradicionales. La Criptografía Ligera o Liviana estudia algoritmos que por sus propiedades matemáticas pueden ejecutarse en plataformas o dispositivos de recursos limitados, como lo son los antes mencionados.

Existen algoritmos livianos de clave privada tipo Block Ciphers⁸, Stream Ciphers⁹ y de clave pública¹⁰ como así también Gestión de Claves, Firma Digital y funciones Hash. Por ejemplo los Block Ciphers creados por la agencia gubernamental NSA¹¹, llamados SIMON y SPECK[7,8] de uso público.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Algoritmos como Simon y Speck, junto con el resto de ellos, deben ser capaces de demostrar su robustez al ser sometidos a ataques criptoanalíticos¹². Es por ello que este proyecto

⁷ 5G: es la llamada 5ta generación de Tecnologías de Telefonía Móvil. Su antecesora, la tecnología 4G aún no del todo difundida en nuestro país.

⁸ Algoritmo de Cifrado por Bloques: algoritmos que dividen el mensaje a cifrar en bloques de n bits y luego proceden al cifrado del bloque.

⁹ Algoritmo de Cifrado en Cadena o Flujo: algoritmos que generan largas secuencias pseudoaleatorias de bits, los cuales uno a uno pueden ser operados con cada bit del mensaje a cifrar.

¹⁰ Algoritmos que utiliza 2 claves, una de ellas es pública y sirve para cifrar el mensaje. La otra permanece secreta y se usa para descifrar el mensaje. También son llamados Algoritmos Asimétricos, por el uso que se hace de sus claves.

¹¹ National Security Agency: Agencia de Seguridad Nacional. Organismo gubernamental de Estados Unidos.

¹² Criptoanálisis: parte de la Criptología que se encarga de analizar, estudiar y desarrollar ataques para el

persigue la profundización en el estudio de las propiedades criptológicas y matemáticas que posibiliten hallar sus vulnerabilidades o debilidades.

3.RESULTADOS OBTENIDOS/ ESPERADOS

El objetivo de este proyecto es abordar y profundizar en el conocimiento de las propiedades criptológicas y de seguridad de Algoritmos Criptográficos Livianos que puedan emplearse en Internet de las Cosas[9] u otros dispositivos semejantes, que así lo requieran por sus limitaciones.

Se realizará un relevamiento, estudio y análisis exhaustivo de los principales algoritmos, que podrían ser usados en IoT, poniendo énfasis en los del tipo Stream Ciphers, pues son los que por sus características podrían emplearse con mayor asiduidad en los dispositivos de IoT y RFID.

Se definirán indicadores utilizando otras experiencias internacionales para evaluar comportamientos y permitir comparaciones.

4.FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigadores pertenece al cuerpo docente de Tecnologías Aplicadas en la Facultad de Ingeniería, el área de la Seguridad Informática, de la Universidad del Salvador.

A fines del año 2017 el alumno *Leonardo Parisi* se ha sumado como colaborador al equipo de investigación. Se espera que en breve más alumnos se incorporen como él.

5. BIBLIOGRAFÍA

[1] <http://www.lanacion.com.ar/1753934-las-zapatillas-con-gps-dan-un-primer-paso-buscando-nuevos-mercados>. Consultada el 1-3-2017.

[2] ISO/IEC 29192. Information technology - Security techniques - Lightweight Cryptography. 2012. <https://www.iso.org>.

[3] Manyika, J.; Chui, M.; Bughin, J.; Dobbs, R.; Bisson, P.; Marrs, A. Disruptive technologies: Advances that will transform life, business, and the global economy. McKinsey Global Institute. 2013.

[4] https://www.ericsson.com/mx/news/2015-11-17-emr-es_254740126_c. Consultada el 1-3-2017.

[5] <https://www.ericsson.com/en/mobility-report/reports/november-2017>. Consultada el 3-2-2018.

[6] http://tn.com.ar/tecno/f5/ces-2016-las-heladeras-del-futuro-conectadas-y-con-multiples-sensores_647274

[7] <http://www.nsa.gov/>. Consultada el 1-3-2017.

[8] <http://eprint.iacr.org/2013/404.pdf>

[9] Masanobu Katagi; Shiho Moriai, Lightweight Cryptography for the Internet of Things; Sony Corporation; 2016.

descubrimiento de los mensajes cifrados o las claves que fueron empleadas.

Entorno de disponibilidad forense para la recolección de datos digitales en HTTP

Mónica D. Tugnarelli ⁽¹⁾, Mauro F. Fornaroli ⁽¹⁾,
Sonia R. Santana ⁽¹⁾, Javier Díaz ⁽²⁾

⁽¹⁾ Facultad de Ciencias de la Administración – Universidad Nacional de Entre Ríos

⁽²⁾ Facultad de Informática – Universidad Nacional de La Plata

e-mail: montug, maufor [@fcad.uner.edu.ar]

Resumen

En este artículo se presentan los avances del análisis de prestaciones y características de dos metodologías de recolección de datos digitales asociadas a eventos de seguridad, la primera llamada Enfoque preventivo-Recolección de datos a priori o *Forensic Readiness* y, la segunda, Enfoque reactivo - Recolección de datos a posteriori de un incidente. Para el desarrollo del trabajo se establecieron etapas y actividades tendientes a la identificación y descripción de puntos de control en protocolos HTTP, la configuración de un entorno de prueba para la captura de datos, procedimientos para la preservación de la evidencia, aspectos de integridad y trazabilidad de la misma y la descripción de resultados comparativos entre ambos enfoques de recolección.

Palabras clave: seguridad, disponibilidad forense, evidencia digital, servidores web, HTTP.

Contexto

El Proyecto de Investigación y Desarrollo PID-UNER 7052 para Director Novel con Asesor, denominado “Análisis de Metodologías de Recolección de datos

digitales” se encuadra en una de las líneas de investigación establecidas como prioritarias para su fomento, "Arquitectura, Sistemas Operativos y Redes", de la carrera Licenciatura en Sistemas de la Facultad de Ciencias de la Administración. Se adecua además, a las prioridades de la Universidad Nacional de Entre Ríos por ser un proyecto aplicado a la investigación sobre Tecnologías de la Información y la Comunicación [1].

Introducción

Si una arquitectura de seguridad informática está correctamente definida debe ofrecer un plan y un conjunto de políticas que describan tanto los servicios de seguridad ofrecidos a los usuarios como los componentes del sistema requeridos para implementar dichos servicios. Estas políticas de seguridad se aplican a los activos de información identificados por su relevancia con los objetivos de la organización, conociendo como se gestionan y cuáles son sus riesgos, con la finalidad de implementar estrategias y mecanismos que aseguren la confidencialidad, la integridad y la disponibilidad de los mismos [2].

En este entorno tecnológico las técnicas y metodologías de forensia informática deben asegurar que se pueda determinar adecuadamente el *qué, quién, cuándo y cómo sucedió* en relación al incidente de seguridad, así como también ocuparse de la correcta preservación y trazabilidad de los datos recolectados. Entonces, el análisis forense digital requiere aplicar métodos científicos, técnicas y herramientas para cumplimentar etapas relacionadas con la identificación, preservación y análisis de la evidencia digital, la cual llegado el caso puede ser considerada legalmente en un proceso judicial [3]. Una etapa destacada es la recolección de esta evidencia y la manera en que se asegura la calidad e integridad de los datos recolectados.

En el marco del proyecto se analizan dos enfoques de recolección de datos:

a) Recolección de datos a priori de un evento de seguridad: también conocido como *Forensic Readiness*. Este enfoque introduce el concepto de resguardar la posible evidencia antes de que ocurra el incidente para cubrir dos objetivos: maximizar la capacidad del entorno para reunir evidencia digital confiable y minimizar el costo forense durante la respuesta a un incidente. [4], [5], [6]. La planificación del entorno de disponibilidad forense incluye, entre otras acciones:

- Definir activos que pueden requerir de pruebas digitales;
- Identificar las fuentes disponibles y los diferentes tipos de pruebas;
- Establecer una forma segura de obtención de pruebas para cumplir con el requisito de admisibilidad legal;

- Establecer una política para el almacenamiento y el manejo seguro de las evidencias;
- Capacitar al personal de modo que todos entiendan su papel en el proceso de pruebas digitales y la sensibilidad jurídica de las mismas.

b) Recolección de datos a posteriori de un evento de seguridad. Este enfoque recupera la evidencia luego de que se haya detectado el incidente de seguridad con el objetivo de realizar un análisis forense para determinar lo ocurrido.

En este proyecto el activo identificado es un servidor web, específicamente analizando información del protocolo HTTP en sus versiones 1.1 y 2 [7], [8], empleando para la ejecución de pruebas y adquisición de datos herramientas de forensia con licenciamiento libre [9], principalmente Kali Linux [10]. Como guía y marco general para las pruebas se consideran las especificaciones de la RFC 3227 - Guidelines for Evidence Collection and Archiving [11], la ISO/IEC 27037- Guidelines for identification, collection, acquisition and preservation of digital evidence [12] y el Open Source Security Testing Methodology Manual (OSSTMM) [13].

Líneas de Investigación, Desarrollo e Innovación

Con este proyecto de investigación se espera conformar una base de conocimiento que aporte al análisis de la forensia informática en relación a metodologías de

recolección de datos digitales y, específicamente a entornos de Forensic Readiness y las buenas prácticas asociadas a su implementación. Se prevé profundizar la línea de investigación con nuevos proyectos y con la firma de convenios de colaboración con otras instituciones de investigación en la temática.

Resultados y Objetivos

El PID 7052 tiene como objetivo principal analizar comparativamente la performance y características de las dos metodologías de recolección de datos en entornos de servidores web.

Como puntos de control HTTP para la recolección de datos, sobre los cuales se realizó un monitoreo activo, la captura de tráfico y recolección de datos, se identificaron los siguientes:

- puertos 80 y 443
- logs del sistema operativo (*/var/log/*): *messages.log*, *auth.log*, *secure*, *utmp/wtmp*,
- logs *httpd*: *error.log* y *access.log*.

Según el enfoque varía el resguardo de información, considerando que para la metodología *Forensic Readiness* cada dato recolectado debe ser asegurado como posible evidencia digital.

En la Figura 1 se muestra la secuencia del procedimiento aplicado para la recolección y aseguramiento de los datos en el Enfoque preventivo [14]:



Figura 1. Secuencia procedimiento enfoque preventivo

En la Figura 2 se presenta la secuencia del procedimiento para el Enfoque reactivo:

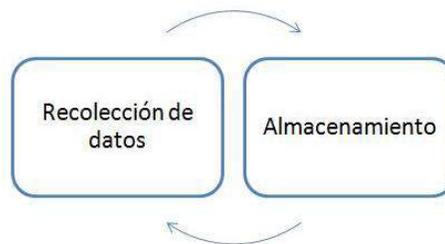


Figura 2. Secuencia procedimiento enfoque reactivo

Como parte de las actividades se realizó la simulación de un ataque de Denegación de Servicios Distribuido (DDoS) a los fines de observar el comportamiento del servidor, el volumen de datos resguardados en logs y para analizar la performance de las metodologías en instancias de recuperación del servicio y resguardo de la evidencia.

Los primeros resultados obtenidos en este proyecto permiten establecer que:

1. La metodología preventiva proporciona un mecanismo activo de anticipación a los incidentes en contraste con las metodologías de respuesta a un incidente de seguridad.

2. En un enfoque preventivo, la integridad de los datos tratados como evidencia digital se puede asegurar con el hash que actúa como una faja digital cumpliendo la misma función que el sellado físico utilizado para resguardar dispositivos comprometidos [15]
3. La trazabilidad de los datos considerados evidencia puede mantenerse implementando un control de versiones, que registra los cambios realizados sobre un archivo o conjunto de archivos a lo largo del tiempo.
4. En relación a la infraestructura mínima necesaria, hay que considerar también el valor temporal que la organización le otorgue a la información recolectada, que puede estar sujeto a políticas de la organización como así también a leyes nacionales e internacionales y/o a restricciones dadas por el propio carácter de los datos. El enfoque preventivo, y a medida que se avanza en la recolección de datos de los activos identificados requiere una infraestructura de alta prestación, cuestión que se abordará en un próximo proyecto de investigación
5. En el enfoque preventivo se maximiza la explotación de potencial evidencia, ya que la misma puede ser preservada y mantenerse no contaminada o dañada por el incidente de seguridad, aspecto que es muy difícil de asegurar en un enfoque reactivo cuando ya un ataque de seguridad se ha producido.

Formación de Recursos Humanos

En este proyecto se forman e inician en actividades de investigación tres docentes de la carrera Licenciatura en Sistemas y un docente en la dirección de proyectos. Asimismo da el marco para la realización de una tesis de maestría correspondiente a la Maestría en Redes de Datos de la Facultad de Informática de la Universidad Nacional de La Plata.

Referencias

- [1] Tugnarelli, M., Fornaroli, M., Santana, S., Jacobo, E., Díaz, J.: Análisis de Metodologías de Recolección de Datos Digitales. In: Libro de Actas Workshop de Investigadores en Ciencias de la Computación 2017, pp. 1000-1004. ISBN 978-987-42-5143-5.
- [2] Incident Management and Response ISACA <http://www.isaca.org/>
- [3] Digital Forensic Research Workshop (DFRWS). <http://www.dfrws.org/>
- [4] TAN, John. (2001). Forensic Readiness. http://isis.poly.edu/kulesh/forensics/forensic_readiness.pdf
- [5] Rowlingson, Robert. A Ten Step for Forensic Readiness. (2004) International Journal of Digital Evidence. Volume 2, Issue 3.
- [6] Poee, A. , Labuschagne, L. A conceptual model for digital forensic readiness (2012) <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=6320452>
- [7] RFC 2616 Hypertext Transfer Protocol - HTTP/1.1 <http://tools.ietf.org/html/rfc2616>

- [8] RFC 7540 Hypertext Transfer Protocol Version 2 (HTTP/2).
<https://tools.ietf.org/html/rfc7540>
- [9] Tugnarelli, M.; Fornaroli, M.; Pacifico, C. Análisis de prestaciones de herramientas de software libre para la recolección a priori de evidencia digital en servidores web. Workshop de Investigadores en Ciencias de la Computación (WICC 2015). ISBN 978-987-633-134-0
- [10] KALI Linux. www.kali.org
- [11] RFC 3227 Guidelines for Evidence Collection and Archiving.
<https://www.ietf.org/rfc/rfc3227.txt>
- [12] Guidelines for identification, collection, acquisition and preservation of digital evidence” ISO/IEC 27037:2012
- [13] Open Source Security Testing Methodology Manual (OSSTM)
<http://www.isecom.org/mirror/OSSTM.M.3.pdf>
- [14] Mónica D. Tugnarelli, Mauro F. Fornaroli, Sonia R. Santana, Eduardo Jacobo, Javier Díaz: Análisis de metodologías de recolección de datos digitales en servidores web. Libro de Actas. XXIII Congreso Argentino de Ciencias de la Computación CACIC 2017. VI Workshop de Seguridad Informática, pp. 1230-1238. ISBN 978-950-34-1539-9.
- [15] Piccirilli, Darío. (2016). Protocolos a aplicar en la forensia informática en el marco de las nuevas tecnologías (pericia – forensia y cibercrimen). Tesis de doctorado. Facultad de Informática. Universidad Nacional de La Plata.
<http://hdl.handle.net/10915/52212>
- [16] U.S. Department of Justice. Electronic Crime Scene Investigation: A Guide for First Responders, Second Edition. <https://www.ncjrs.gov/pdffiles1/nij/219941.pdf>, last accessed 2017/08/30.
- [17] Forte, D: Principles of digital evidence Collection. Elsevier, Network Security, Volume 2003, Issue 12, 6-7 (2003).
- [18] Caracciolo Claudio, Rodriguez Marcelo, Sallis Ezequiel. (2010). *Ethical Hacking - un enfoque metodológico para profesionales*
- [19] Computer Forensics. (2008) Volume 56, Number 1. U.S. Department of Justice
- [20] FBI Cyber Crime.
<http://www.fbi.gov/about-us/investigate/cyber>
- [21] Jarrett, Marshall. Bailie, Michael W. Electronic Evidence in Criminal Investigations. Computer Crime and Intellectual Property Section
- [22] del Peso Navarro, Emilio y colaboradores. (2001) *Peritajes Informáticos*. Editorial Díaz de Santos
- [23] Noblett, M., Pollitt, M., Presley, L. (2000). *Recovering and Examining Computer Forensic Evidence*. Forensic Science Communications. Volume 2, Number 4. U.S. Department of Justice. Federal Bureau of Investigation (FBI)
- [24] Digital Evidence and Computer Crime. Forensic Science, Computers and Internet. Third Edition (2011). Eoghan Casey. Elsevier Inc.

Evaluación de la Seguridad en Sistemas Informáticos

Aristides Dasso, Ana Funes, Daniel Riesco, Germán Montejano

SEG / Departamento de Informática / Facultad de Ciencias Físico-
Matemáticas y Naturales / Universidad Nacional de San Luis
Ejército de los Andes 950, D5700HHW San Luis, Argentina
+54 (0) 266 4520300, ext. 2126
{arisdas, afunes, driesco, gmonte}@unsl.edu.ar

RESUMEN

En este trabajo presentamos los objetivos, lineamientos y resultados de una línea de investigación sobre la creación de modelos de evaluación de seguridad informática en organizaciones. En el ámbito del desarrollo de modelos de evaluación de sistemas complejos, esta investigación tiene como objetivo la creación, puesta a punto y aplicación de modelos que permitan obtener indicadores del nivel de madurez alcanzado en la seguridad de un sistema informático.

En general, el método empleado para el desarrollo de dichos modelos de evaluación es el método LSP. Asimismo, hemos tomado como referencia para la creación de partes de los modelos, estándares reconocidos como son la norma ISO 27000, o recomendaciones dadas por organismos internacionales, como la OWASP, que sirven de guía a las organizaciones para formular e implementar estrategias para la seguridad del software.

Palabras clave: Seguridad de Sistemas Informáticos. Evaluación de la Seguridad de Sistemas Informáticos. Métodos de Evaluación. Logic Scoring of Preference (LSP). SAMM. Software Assurance Maturity Model.

CONTEXTO

Este trabajo de investigación se viene llevando a cabo dentro del SEG (Software Engineering Group), en el ámbito de la

Universidad Nacional de San Luis y se encuentra enmarcado dentro de una de las líneas de investigación del Proyecto de Incentivos código 22/F222 “Aspectos de alta sensibilidad en el ejercicio de la profesión de Ingeniero de Software” (Director: Daniel Riesco, Co-Director: Roberto Uzal. Acreditado con evaluación externa. Financiamiento: Universidad Nacional de San Luis).

INTRODUCCIÓN

Los modelos de evaluación de sistemas de seguridad permiten conocer, dentro de una organización, la madurez con la cual la organización lleva adelante sus políticas, actividades, usa sus herramientas y métodos, etc., en pos de su seguridad.

Si bien la construcción de modelos de evaluación de sistemas complejos, entre los que se encuentran los sistemas de seguridad informática, constituye una necesidad importante, no es una tarea sencilla. Múltiples aspectos deben ser considerados en esta tarea; teniendo en cuenta no solo los aspectos físicos, tales como las instalaciones y sus políticas de acceso, sino también medidas de seguridad del software, tales como firewalls, permisos, codificación en línea, etc.

Por lo tanto, para una organización preocupada en su seguridad informática, resulta necesario contar con estándares así como herramientas apropiadas para evaluar el grado de adecuación con dichos estándares.

En este sentido, existen en la literatura

múltiples propuestas. Por un lado, se encuentran diversas publicaciones de modelos de madurez de la seguridad, como por ejemplo en el trabajo de M. F. Saleh [18] o el trabajo de Karokola et al. [15], estos últimos, por ejemplo, proponen un modelo para servicios de e-gov; por otro lado, el modelo propuesto por Sjelín et al. [24], ayuda en la implementación de programas de seguridad en áreas comunitarias; la propuesta de S. Almuhammadi et al. [1], complementa el NIST Cybersecurity Framework for Critical Infrastructure; o la propuesta de S. Monteiro et al. [19], quienes presentan una metodología para evaluar la madurez de una organización con respecto a su seguridad. En [4] los autores también presentan una propuesta de modelo basado en los estándares dados en la ISO/IEC 27002.

Asimismo, existen múltiples organismos gubernamentales, educativos, así como instituciones internacionales, que certifican seguridad o que proponen estándares o pruebas para distintos aspectos de la seguridad; entre ellos el CERT, Division del Software Engineering Institute (SEI) [3] de la Carnegie Mellon University (CMU), creado en 1988 como respuesta al Morris worm y el National Institute of Standards and Technology (NIST) of USA que ha desarrollado SCAP (Security Content Automation Protocol) [20], el cual combina varios estándares, esquemas, etc. que pueden ser usados para automatizar diversas medidas de seguridad. La IEEE [17], [23] también tiene sus propios estándares e iniciativas en el campo de la seguridad informática, en particular el IEEE Centre for Secure Design (CSD). [16].

Dentro de los organismos internacionales, también se encuentra la Open Web Application Security Project (OWASP), una organización mundial sin fines de lucro, focalizada en mejorar la seguridad del software, la cual cuenta con varios proyectos relacionados con la seguridad informática, entre ellos el proyecto SAMM (Software Assurance Maturity Model) [22]. SAMM es un framework abierto creado para ayudar a las organizaciones en varios aspectos de la

seguridad, tales como la evaluación, el desarrollo de programas de seguridad, programas de mejoras, y definición y medición de actividades relacionadas con la seguridad.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La línea de investigación en la que se enmarca el trabajo presentado, es parte de una línea de investigación sobre el tema de la construcción de modelos de evaluación de sistemas complejos y que viene desarrollándose desde hace tiempo en el ámbito del SEG (Software Engineering Group), donde se han obtenido resultados plasmados en diversas publicaciones (ver por ejemplo [12][13][4] [14][2][5][6]).

En esta instancia, nos encontramos abocados a estudiar la propuesta de la OWASP, particularmente en su proyecto SAMM [22], con el objetivo de proponer un nuevo modelo de evaluación de la seguridad, que si bien, se basa en SAMM, apunta a mejorar la forma de evaluación. Más concretamente, como parte de las herramientas de OWASP, se encuentra el SAMM Assessment Toolbox v1.5 [21]. Se trata de un template que sirve de guía al equipo evaluador durante las entrevistas llevadas a cabo dentro de la organización a ser evaluada. Usando este Toolbox, dicho equipo cuenta con una guía sobre qué aspectos debe abordar durante las entrevistas y, además, puede asignar puntajes a cada pregunta. La herramienta, de forma automática, al finalizar la evaluación devuelve un puntaje global que es interpretado como un indicador de la madurez alcanzada dentro de la organización en su seguridad informática. Para esto, la herramienta adopta un simple método de puntaje aditivo. Sin embargo, creemos que la evaluación de la seguridad informática involucra una serie de decisiones del tipo y/o que deben considerarse en su momento.

En este sentido, el método LSP (Logic Score of Preference) [7][8][9][10][11] viene a cubrir esta necesidad ya que se trata de un método que permite la creación de modelos

de evaluación, especialmente útil para sistemas complejos, en los cuales un número importante de decisiones deben ser consideradas.

RESULTADOS Y OBJETIVOS

Actualmente, nos encontramos desarrollando el modelo de evaluación, para lo cual hemos adoptado el método LSP.

En primer lugar, ya hemos establecido un conjunto de requerimientos, bajo la forma de una estructura jerárquica, llamado árbol de

requerimientos en LSP. Estos requerimientos provienen directamente de cada nivel del SAMM.

Como se puede ver en la Tabla 1, SAMM se encuentra organizado en cuatro funciones principales del negocio (Governance, Construction, Verification and Operations). Cada función del negocio comprende a su vez tres prácticas de seguridad, y cada una de esas tres prácticas cuenta con tres niveles de seguridad u objetivos. Cada uno de esos objetivos se encuentra definido por métricas y actividades específicas.

Tabla 1. Los tres primeros niveles de SAMM

Business Functions	Security Practices	Objectives
1. Governance	1.1. Strategy & Metrics	1.1.1. SM1: Establish a unified strategic roadmap for software security within the organization.
		1.1.2. SM2: Measure relative value of data and software assets and choose risk tolerance.
		1.1.3. SM3: Align security expenditure with relevant business indicators and asset value.
	1.2. Policy & Compliance	1.2.1. PC1: Understand relevant governance and compliance drivers to the organization
		1.2.2. PC2: Establish security and compliance baseline and understand per-project risks.
		1.2.3. PC3: Require compliance and measure projects against organization-wide policies and standards.
	1.3. Education & Guidance	1.3.1. EG1: Offer development staff access to resources around the topics of secure programming and deployment.
		1.3.2. EG2: Educate all personnel in the software lifecycle with role-specific guidance on secure development.
		1.3.3. EG3: Mandate comprehensive security training and certify personnel for baseline knowledge.
2. Construction	2.1. Threat Assessment	2.1.1. TA1: Identify and understand high-level threats to the organization and individual projects.
		2.1.2. TA2: Increase accuracy of threat assessment and improve granularity of per-project understanding.
		2.1.3. TA3: Concretely align compensating controls to each threat against internal and third-party software.
	2.2. Security Requirements	2.2.1. SR1: Consider security explicitly during the software requirements process.
		2.2.2. SR2: Increase granularity of security requirements derived from business logic and known risks.
		2.2.3. SR3: Mandate security requirements process for all software projects and third-party dependencies.
	2.3. Secure Architecture	2.3.1. SA1: Insert consideration of proactive security guidance into the software design process.
		2.3.2. SA2: Direct the software design process toward known secure services and secure-by-default designs.
		2.3.3. SA3: Formally control the software design process and validate utilization of secure components.
3. Verification	3.1. Design Review	3.1.1. DR1: Support ad-hoc reviews of software design to ensure baseline mitigations for known risks.
		3.1.2. DR2: Offer assessment services to review software design against comprehensive best practices for security.
		3.1.3. DR3: Require assessments and validate artifacts to develop detailed understanding of protection mechanisms.
	3.2. Implementation Review	3.2.1. IR1: Opportunistically find basic code-level vulnerabilities and other high-risk security issues.
		3.2.2. IR2: Make implementation review during development more accurate and efficient through automation.
		3.2.3. IR3: Mandate comprehensive implementation review process to discover language-level and application-specific risks.
	3.3. Security Testing	3.3.1. ST1: Establish process to perform basic security tests based on implementation and software requirements.
		3.3.2. ST2: Make security testing during development more complete and efficient through automation.
		3.3.3. ST3: Require application of specific security testing to ensure baseline security before deployment.
4. Operations	4.1. Issue Management	4.1.1. IM1: Understand high-level plan for responding to issue reports or incidents.
		4.1.2. IM2: Elaborate expectations for response process to improve consistency and communications.
		4.1.3. IM3: Improve analysis and data gathering within response process for feedback into proactive planning.
	4.2. Environment Hardening	4.2.1. EH1: Understand baseline operational environment for applications and software components.
		4.2.2. EH2: Improve confidence in application operations by hardening the operating environment.
		4.2.3. EH3: Validate application health and status of operational environment against known best practices.
	4.3. Operational Enablement	4.3.1. OE1: Enable communications between development teams and operators for critical security-relevant data.
		4.3.2. OE2: Improve expectations for continuous secure operations through provision of detailed procedures.
		4.3.3. OE3: Mandate communication of security information and validate artifacts for completeness.

El cuarto nivel del árbol, que no se muestra en la tabla por razones de espacio, contiene un número de preguntas que son formuladas durante las entrevistas, mientras que en el quinto, se acomodan las hojas del árbol, llamadas variables de performance en LSP, que corresponden a las guías de entrevista dados en el SAMM Assessment Toolbox v.1.5.

De acuerdo al método LSP, durante el proceso de evaluación, cada variable de performance debe ser transformada en una preferencia elemental, un valor entre 0 y 100, que representa el grado de satisfacción alcanzado en ese ítem. Para esto, en ocasiones resulta necesario definir algunas funciones, llamadas criterios elementales en LSP. En nuestro caso, dado que cada preferencia elemental es asignada en forma directa por el evaluador no hay necesidad de definir ningún criterio elemental.

La última etapa de la construcción del modelo es la creación y puesta a punto de la estructura de agregación o función de criterio LSP. La misma se encuentra aún en desarrollo ya que el proceso de puesta a punto implica la evaluación de varios sistemas para ver cómo se comporta el modelo y su adaptación para ir ajustándolo.

Dicha estructura de agregación se construye a partir de las hojas del árbol de requerimientos, por medio de un proceso iterativo que comienza por agregar grupos de preferencias elementales, por medio de operadores provisto por el método, y generando un número de preferencias agregadas. Este es un proceso bottom up que se repite sobre las preferencias agregadas hasta que se obtiene una única preferencia global. Este resultado global, que es un valor en el intervalo $[0..100]$, representa el porcentaje de cumplimiento con respecto a los requisitos de seguridad establecidos en el árbol (y en SAMM) y es un indicador, en nuestro, caso del nivel de madurez alcanzado por la organización bajo evaluación.

Como parte del trabajo futuro, esperamos, en una etapa siguiente, finalizar con la calibración de la estructura de agregación y

poder aplicar el modelo a la evaluación de casos reales.

FORMACIÓN DE RECURSOS HUMANOS

Dentro del SEG (Software Engineering Group), en el ámbito de la Universidad Nacional de San Luis, en el que se realiza el Proyecto de Incentivos código 22/F222 “Aspectos de alta sensibilidad en el ejercicio de la profesión de Ingeniero de Software”, se han llevado a cabo numerosas tesis de grado y de posgrado.

Entre otros, nos hemos concentrado en la evaluación de sitios de gobierno electrónico lo que ha dado como resultado una tesis de maestría en 2010 [2], mientras que hay otra en etapa de finalización sobre modelos de evaluación de la accesibilidad web [14]. La construcción del modelo aquí expuesto, también, tiene como objetivo ser motivo de tesis, como lo han sido la construcción de otras herramientas en el ámbito del proyecto.

REFERENCIAS

- [1] Almuhammadi, Sultan, Alsaleh, Majeed. “INFORMATION SECURITY MATURITY MODEL FOR NIST CYBER SECURITY FRAMEWORK”. Computer Science & Information Technology, Sixth International Conference on Information Technology Convergence and Services (ITCS 2017). Sydney, Australia, February 25–26, 2017.
- [2] Castro, M. Dasso, A., Funes, A. “Modelo de Evaluación para Sitios de Gobierno Electrónico”. 38 JAIIO/SIE 2009, Simposio de Informática en el Estado 2009, Mar del Plata, Argentina, August 26-28, 2009.
- [3] CERT Division of the Software Engineering Institute (SEI). Carnegie Mellon University (CMU). <http://www.cert.org/> (Retrieved March 2015)
- [4] Dasso, Aristides, Funes, Ana, Montejano, Germán, Riesco, Daniel, Uzal, Roberto, Debnath, Narayan; “Model Based Evaluation of Cybersecurity Implementations”. ITNG 2016. Las Vegas, Nevada, USA, 11-13 abril 2016. In S. Latifi (ed.), Information Technology New Generations, Advances in Intelligent Systems and Computing 448. DOI: 10.1007/978-3-319-32467-8_28. Springer International Publishing, Switzerland 2016.
- [5] Debnath, N., Dasso, A., Funes, A., Montejano, G., Riesco, D., Uzal, R. “The LSP Method Applied to Human Resources Evaluation and Selection”, Journal of Computer Science and Information Management, Publication of the Association of Management/International Association of Management, Volume 3, Number 2, 2000, ISBN 1525-4372, pp.1-12.
- [6] Debnath, N., Dasso, A., Funes, A., Uzal, R., Paganini, J. “E-

- government Services Offerings Evaluation Using Continuous Logic". 2007 ACS/IEEE International Conference on Computer Systems and Applications, AICCSA '2007, Amman, Jordan. Sponsored by IEEE Computer Society, Arab Computer Society, and Philadelphia University, Jordan. May 13-16, 2007.
- [7] Dujmovic, J. J. and Elnicki, R., "A DMS Cost/Benefit Decision Model: Mathematical Models for Data management System Evaluation, Comparison, and Selection", National Bureau of Standards, Washington, D.C., No. NBS-GCR-82-374, NTIS No. PB82-170150 (155 pages), 1982.
- [8] Dujmovic, J. J.: "A Method for Evaluation and Selection of Complex Hardware and Software Systems", The 22nd International Conference for the Resource Management and Performance Evaluation of Enterprise Computing Systems. CMG96 Proceedings, vol. 1, pp.368-378, 1996.
- [9] Dujmovic, J. J.: "Quantitative Evaluation of Software", Proceedings of the IASTED International Conference on Software Engineering, edited by M.H. Hamza, pp. 3-7, IASTED/Acta Press, 1997.
- [10] Dujmovic, J. J.; Bayucan, A.: "Evaluation and Comparison of Windowed environments", Proceedings of the IASTED Interna Conference Software Engineering (SE'97), pp 102-105, 1997.
- [11] Dujmovic, Jozo J.: "Continuous Preference Logic for System Evaluation", IEEE Transactions on Fuzzy Systems, Vol. 15, N° 6, December 2007
- [12] Funes, A., Dasso, A., "Web Application Frameworks Evaluation", CONAISI 2014, 13 y 14 de noviembre de 2014, San Luis, Argentina. pp. 1063-1070. ISSN: 2346-9927.
- [13] Funes, A., Dasso, A., Salgado, C., Peralta, M., "UML Tool Evaluation Requirements", Argentine Symposium on Information Systems ASIS 2005. Rosario, Argentina. September 29-30, 2005.
- [14] Gallardo, Cecilia, Funes, Ana. "Un Modelo para la Evaluación de la Calidad de la Accesibilidad al Contenido Web", CONAISI 2015, 19 y 20 de Noviembre de 2015, Buenos Aires, Argentina, ISBN: 978-987-1896-47-9.
- [15] Geoffrey Karokola, Stewart Kowalski and Louise Yngström. "Towards An Information Security Maturity Model for Secure e-Government Services: A Stakeholders View". Proceedings of the Fifth International Symposium on Human Aspects of Information, Security & Assurance (HAISA 2011). Editors: Steven Furnell, Nathan Clarke. London, United Kingdom 7-8 July 2011. Publisher: Lulu.com - 2011
- [16] IEEE Computer Society Center for Secure Design. <http://cybersecurity.ieee.org/center-for-secure-design.html> (Retrieved March 2015)
- [17] IEEE Cyber Security Initiative. <http://cybersecurity.ieee.org/about.html> (Retrieved March 2015)
- [18] Malik F. Saleh. "Information Security Maturity Model". INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND SECURITY (IJCSS). Edited by DR. NABEEL TAHIR. Volume 5, Issue 3, August 2011. Publishing Date: July / August 2011. ISSN (Online): 1985 -1553
- [19] Monteiro, S. "Information Security Maturity Level: A Fast Assessment Methodology". In proceedings of Ambient Intelligence-Software and Applications -8th International Symposium on Ambient Intelligence (ISAmI 2017). 21-23 June, Porto, Portugal. J.F. De Paz, V. Julián, G. Villarrubia, G. Marreiros, P. Novais (Eds.). ISBN 978-3-319-61118-1
- [20] National Institute of Standards and Technology (NIST), Security Content Automation Protocol (SCAP) <http://scap.nist.gov/index.html> (Retrieved March 2015)
- [21] OWASP. SAMM_Assessment_Toolbox_v1.5_FINAL.xlsx. Retrieved 28/11/2017 from: https://www.owasp.org/index.php/OWASP_SAMM_Project
- [22] OWASP. Software Assurance Maturity Model. A guide to building security into software development. Version 1.5. OWASP The Open Web Application Security Project. Retrieved 28/11/2017 from: https://www.owasp.org/index.php/OWASP_SAMM_Project
- [23] Rozenfeld, M.: "IEEE Standards on Cyber Security". THE INSTITUTE. IEEE. Volume 39, Issue1, March 2015.
- [24] Sjelín N., White G. "The Community Cyber Security Maturity Model". In: Clark R., Hakim S. (eds) Cyber-Physical Security. Protecting Critical Infrastructure, vol 3. Springer, Cham.

Protección de activos vinculados con la información: preparación para la Ciberdefensa

Alejandro Moresi¹ y Antonio Castro Lechtaler¹

¹ Universidad de la Defensa, Escuela de Guerra Conjunta de las Fuerzas Armadas, Proyecto: Observatorio de CiberDefensa - OBSCIBER, Buenos Aires, Argentina, C1426.

alejandro.moresi@conjunta.undef.edu.ar; antonio.castrolechtaler@gmail.com

RESUMEN

Como consecuencia de diversos estudios realizados en distintas instituciones, se define la necesidad de crear un foro de información y conocimiento de la problemática de un nuevo ambiente operacional, que es la Ciberespacio. Éste fue muy bien caracterizado por su virtualidad, y por el Manifiesto de John Perry Barlow con la Declaración de Independencia del Ciberespacio¹.

Hasta nuestro días, éste ámbito de la vida es exclusivo y único del ser humano. Crece y se multiplica a cada instante de la mano de las llamadas: “Tecnologías de la Información y las Comunicaciones” (TIC), y se extiende al mundo de la movilidad, con el continuo crecimiento del uso de diversos equipos que utilizan las transmisiones inalámbricas, los que pueden ir desde un simple teléfono móvil a equipos más sofisticados.

Por otra parte, el acceso de estos equipos al uso de anchos de banda cada día mayores, permiten millones de transacciones por segundo y movimientos de información entre variados puntos del planeta, en una cantidad hasta hace muy poco impensada.

Ello ha provocado un cambio en los hábitos de la humanidad que impacta de manera radical en todos ámbitos de la vida.

La Universidad de la Defensa Nacional (UNDEF), a través del Programa UNDEFI, se encuentra financiando un Proyecto de un Observatorio de Ciberdefensa a través de la Escuela Superior de Guerra Conjunta de las

Fuerzas Armadas (ESGC), la que para cumplir su cometido, inicialmente se ha asociado mediante convenios ya firmados con la Facultad de Informática de la Universidad de La Plata y con acuerdos con la Subsecretaría de Ciberdefensa del Ministerio de Defensa y el Centro de Estudios General Mosconi de la Escuela Superior Técnica del Ejército, también perteneciente a la UNDEF. También se espera en el futuro, incorporar otros organismos privados y estatales como parte de este esfuerzo.

Este proyecto para desarrollar y poner en funcionamiento un Observatorio de Ciberdefensa fue recientemente sometido a un concurso público con jurados externos, dentro de un Programa de Investigación de la citada universidad denominado UNDEFI.

En dicho marco de referencia, y luego de esa evaluación, al mismo se le otorgó un subsidio especial [1] de \$ 100.000,00 para comenzar su ejecución durante el año 2018.

La idea del mismo es brindar al público relacionado e interesado en conocimientos de nivel político, estratégico y tecnológico de alto nivel, tanto en el plano internacional como nacional, información actualizada acerca de problemáticas tales como: Ciberdefensa; Ciberseguridad; Cibercrimen y Ciberterrorismo, entre otros.

Palabras Clave:

Ciberdefensa - Ciberseguridad - Cibercrimen - Ciberterrorismo.

¹https://danwin1210.me/uploads/%20_%CE%90%C5%8B%E2%82%A3%C3%B8%E1%B4%9A%CE%94%E2%80%A0%E1%B8%AF%C3%B8%CE%B7%20%C3%8E%C2%A7%20%E2%92%B6%C4%BF%C2%A1%E2%93%8B%CE%9E/manifiesto_de_john_perry_barlow-1.pdf

CONTEXTO

El Observatorio de CiberDefensa fue creado por la Escuela Superior de Guerra Conjunta de las Fuerzas Armadas, dependiente de la Universidad de la Defensa (UNDEF). El mismo se encuentra instalado en dicha Escuela Superior y está orientado a satisfacer distintas necesidades de la Defensa Nacional en general y del Ministerio correspondiente en particular.

Tiene por objeto el desarrollo de un programa de investigación, extensión y formación de recursos humanos instituido en el año 2017, conformado por un conjunto de entidades educativas que fueron convocadas por la Escuela Superior de Guerra Conjunta de las Fuerzas Armadas dependiente de la Universidad de la Defensa (UNDEF).

En el futuro funcionará también como un órgano consultivo.

Según ISACA (*Information Systems Audit and Control Association*)², se define la CiberDefensa como “*protección de activos de información, a través del tratamiento de amenazas que ponen en riesgo la información que es procesada, almacenada y transportada por los sistemas de información que se encuentran interconectados*”. [2]

La idea es poner en el conocimiento de especialistas en Sistemas de Seguridad Informática, Criptografía, CiberDefensa y Ciberseguridad, etc., mediante informes trimestrales, cuál es el estado del arte en la materia, desde una perspectiva estratégica y técnica.

La información se obtendrá por medios propios y mediante el uso de un motor de búsqueda específico para el tema, para lo cual se ha tomado contacto y se han firmado convenios con otras universidades nacionales. El motor de búsqueda está configurado para un rastreo de carácter global sobre portales y redes específicas, así como en medios periodísticos y otros similares, desde los cuales se pueda capturar información relacionada con sistemas cibernéticos, los cuales se clasificarán en:

- Tecnologías cibernéticas
- Ataques e incidentes cibernéticos
- Estrategias en el Ciberespacio
- Ciber Forensia

A partir de la información recogida, el equipo de investigación realizará un análisis de la misma y sacará las conclusiones de cómo los hechos censados pueden actuar de forma sinérgica con elementos propios del sistema de ciberdefensa.

Ello permitirá advertir acerca de nuevas capacidades detectadas en el ambiente ciberespacial y establecer la manera en que el mundo orientará su accionar en el espacio cibernético, observar las tácticas y acciones desarrolladas a través del ciberespacio y realizar un análisis tecnológico acerca de cómo los sistemas pueden haber sido vulnerados y/o atacados, cuáles fueron las contramedidas y cómo las mismas han sido detectadas.

Paralelamente, se establecerán las estrategias y características para la preparación y adiestramiento de recursos humanos propios. La información será diseminada a través de una página web creada al efecto, la que contendrá foros de tratamiento de los temas específicos, además del envío personalizado a entidades, centros estratégicos empresas, medios y personal relacionado o interesado en la materia.

1. INTRODUCCIÓN.

El ciberespacio [3] constituye un ámbito virtual, nuevo e intangible creado por los medios informáticos a partir de los diversos modos de conectarse, los cuales si bien constituyen una infraestructura de comunicaciones y sistemas informáticos, desde el punto de vista físico, los conceptos, ideas y acciones que en ese ámbito circulan son estrictamente procesos abstractos propios de la virtualidad.

Poseen la capacidad de dañar física, intelectual y moralmente, ello los convierte en un ámbito de construcción de poder, en este caso el poder del conocimiento y el convencimiento, de allí la necesidad, por parte de quienes ejercen el poder, de

² Asociación de Auditoría y Control sobre los Sistemas de Información.

establecer medidas para su vigilancia, control y explotación.

El elemento de acción en este campo es el software, concepto inherente a la forma de accionar con los medios informáticos y esencia que caracterizada por diferentes nombres: sistemas operativos, navegadores, aplicaciones de trabajo profesional, procesadores de textos, planillas de cálculo, bases de datos, programas de diseño, CAD-CAM, virus, gusanos, troyanos, malware, spyware, ransomware, antivirus, firewall, entre otros.

Todos ellos actúan en el ciberespacio proyectando las capacidades de quien desea ejercer acciones, tales como transmitir una información, generar un conocimiento, promover un espacio de reflexión, diversión o esparcimiento, o quien pretende ejercer el poder desde este nuevo ambiente operacional. La República Argentina ha iniciado su accionar en el tema de seguridad informática, a través de carreras de postgrado en diferentes universidades y centros de investigación. Un segundo peldaño, en la cuestión ciberespacial, se constituye desde el aspecto de la Defensa en el año 2014, con la creación del Comando Conjunto de Ciberdefensa, Resolución MD N° 343/14 de fecha 14 de mayo de 2014, cuya misión es: *“Ejercer la Conducción de las Operaciones de Ciberdefensa en forma permanente a los efectos de garantizar las Operaciones Militares del Instrumento Militar de la Defensa Nacional en cumplimiento de su misión principal y de acuerdo a los lineamientos establecidos en el Planeamiento Estratégico Militar”*.

La seguridad de la información es un aspecto tratado por el hombre desde el principio de los tiempos, se conceptualizó la idea de información como elemento de poder de los estados, en los 80/90 a partir de diferentes estudios. La información comenzó a ser considerada un objeto esencial al poder (más allá de que siempre lo fue).

Fue en el siglo XXI, donde la información y la capacidad de incidir sobre el pensamiento de las sociedades pasaron a tener una preeminencia estratégica superior, que se manifiesta con mayor fuerza e intensidad,

durante los procesos de crisis y conflicto. Esto ocurre, tanto en el campo político, como en el militar, el económico o el social.

Esta capacidad, y las características mismas de tratarse de un ambiente completamente nuevo y casi libre de restricciones, permiten que tanto los Estados, como organizaciones lícitas e ilícitas, puedan explotar su empleo para beneficio de la humanidad o para cometer sabotaje, espionaje y otras acciones delictivas.

La característica de intangibilidad, velocidad y capacidad de escalar las acciones que se llevan a cabo en el ciberespacio, constituyen un factor de vulnerabilidad y desestabilización social, que no solo es difícil de prevenir, sino de controlar con el evento en ejecución.

Analizando lo que establece el artículo 2° de la Ley de Defensa Nacional -Ley N° 23.554- es esencial a la Defensa poseer la capacidad de ejecutar medidas contundentes tendientes a impedir la supremacía informática de cualquier posible enemigo de la sociedad argentina [4].

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO.

Las líneas de Investigación y desarrollo son las siguientes:

- 2.1. Desarrollar un servicio de excelencia, con información en el estado del arte, respecto de la problemática de control y negación del ciberespacio, métodos, tecnologías y estrategias de aplicación.
- 2.2. Realizar investigaciones sobre el marco de la ciberdefensa de forma general y aplicada.
- 2.3. Establecer y comprobar los elementos necesarios para tener sistemas de información fiable.
- 2.4. Estudiar y sugerir propuestas orientadas a la toma de decisiones sobre las políticas en aspectos relativos a la CiberDefensa y Ciberseguridad.
- 2.5. Crear un Reservorio de Documentación que estimule el estudio y la investigación.
- 2.6. Realizar y fomentar encuentros entre profesionales y expertos sobre los resultados

obtenidos en investigaciones similares e intercambiar experiencias comunes.

2.7. Elaborar planes de formación sobre temas de Ciberseguridad y CiberDefensa a nivel profesional y para la ciudadanía en general.

2.8 Medir el estado actual y la evolución del bienestar y calidad de vida en base a la mayor concienciación en materia de CiberDefensa y CiberSeguridad.

2.9. Conformar un grupo de analistas relacionados con el área, a los efectos de elaborar estudios y conclusiones sobre los datos obtenidos, que permitan,

- Actuar de forma sinérgica con elementos propios del sistema de ciberdefensa.
- Advertir acerca de nuevas capacidades detectadas en el ambiente ciberespacial.
- Observar el estado del arte en el nivel mundial y su orientación en el cibernético.
- Conocer y difundir tácticas y acciones desarrolladas a través del ciberespacio.

2.10. Investigar los distintos trabajos existentes sobre la seguridad que ofrecen los diferentes tipos de Redes de Comunicaciones [5] para su uso seguro y a los fines de mejorar los aspectos que hacen a la toma de decisiones acerca de su implementación en distintos ambientes.

3. RESULTADOS OBTENIDOS / ESPERADOS.

Como este proyecto es de reciente constitución, a la fecha no se pueden mostrar aún, resultados obtenidos.

Se esperan obtener los siguientes resultados:

3.1. Generar una base de datos sobre CiberDefensa, Ciberseguridad, su estado del arte y de los sistemas asociados a estas materias.

3.2. Generación de indicadores objetivos que permitan la correlación y comparación evolutiva de variables relacionadas con esta temática.

3.3. Publicar y difundir los materiales generados a través de los diferentes estudios e investigaciones.

3.4. Elaborar un reporte anual de CiberDefensa, en el cual se recoja un diagnóstico que permita conocer la situación y cuáles son las herramientas disponibles para la evaluación y la medida de los diferentes fenómenos que intervienen en la percepción de la Defensa.

3.4 Asesorar técnicamente a Fundaciones, Administraciones, Organizaciones, Sociedades, Instituciones y Universidades en materia de CiberDefensa y Ciberseguridad así como seguir los proyectos de cooperación exterior.

3.5. Sistematizar, analizar y procesar datos, para obtener información útil, generar reportes y opiniones de expertos a diferentes usuarios, así como métricas que serán el objeto de este observatorio.

3.6. Creación de materiales pedagógicos dirigidos a los jóvenes en particular y a la ciudadanía en general, en el ámbito de la CiberDefensa.

3.7 Plasmar en un Foro Anual presencial las investigaciones y conclusiones que se van manejando en el ámbito de la CiberDefensa y Ciberseguridad.

4. FORMACIÓN DE RECURSOS HUMANOS.

Desde el año 2017 en este grupo trabajan investigadores formados y categorizados, investigadores en formación, y alumnos de las carreras de grado y posgrado vinculadas con los temas que hacen a la Ciberdefensa y a la Ciberseguridad en el ámbito de la Defensa Nacional, fundamentalmente aquellos pertenecientes a las carreras que se desarrollan en la Escuela Superior de Guerra Conjunta de las Fuerzas Armadas; como así también en la Escuela Superior de Guerra (Ejercito), Escuela Superior de Guerra Naval, Escuela Superior de Guerra Aérea y Escuela Superior Técnica del Ejercito.

Durante este año 2018, se sumarán al proyecto nuevos investigadores formados y en formación y alumnos de las distintas Unidades Académicas antes mencionadas.

Existe también la posibilidad de que algunos alumnos de las instituciones mencionadas,

realizaran monografías sobre alguno de los temas que aborda la presente línea de investigación, ya sea dentro de los planes de estudio, o como trabajos de fin de carrera. Atendiendo a la responsabilidad ética y social que compete a la actividad científica y tecnológica, el Grupo Integrante de este Proyecto de Investigación, ya sea durante su ejecución o por la aplicación de los resultados obtenidos, desea expresar su compromiso en abstenerse de realizar cualquier actividad personal o colectiva que pudiera afectar los derechos humanos, o ser causa de un eventual daño al medio ambiente, a los animales y/o a las generaciones futuras.

5. BIBLIOGRAFÍA.

[1] Resolución Rectoral UNDEF N° 282/2017 de fecha 27 de octubre de 2017. Expediente N° 606/2016.

[2] <https://www.isaca.org/pages/default.aspx>

[3] Julio Gerardo Lucero. La Dimensión Desconocida. El ciberespacio como nuevo ámbito operacional. CEFADIGITAL. Repositorio Digital del Centro Educativo de las Fuerzas Armadas. <http://www.cefadigital.edu.ar/bitstream/123456789/225/1/VC%2012-2015%20LUCERO.pdf>

[4] Torres Soriano, Manuel. Hackeando a la Democracia: operaciones de influencia en el ciberespacio. Instituto Español de Estudios Estratégicos. 2017.

[5] Corletti Estrada, Alejandro. Seguridad en Redes. Learning Consulting S. L. Madrid. 2016.

Confidencialidad de las comunicaciones en Sistemas Móviles

Castro Lechtaler, Antonio^{1,2}; Cipriano, Marcelo^{1,3}; García, Edith¹,
Liporace, Julio¹; Maiorano, Ariel¹; Malvacio, Eduardo¹; Tapia, Néstor¹;

¹Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática.
Escuela Superior Técnica, Facultad del Ejército. Universidad de la Defensa Nacional UNDEF.

²CISTIC/FCE - Universidad de Buenos Aires.

³Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes UNQ.

acastro@est.iue.edu.ar , marcelocipriano@est.iue.edu.ar,
{edithxgarcia; jcliporace; maiorano; edumalvacio; tapianestor87}@gmail.com

RESUMEN.

El objetivo de este proyecto es el diseño y desarrollo de un **Algoritmo de Cifrado** que permita dotar de confidencialidad a la información transmitida en *Sistemas Móviles*. Un algoritmo así requiere satisfacer requerimientos de robustez, compactibilidad y velocidad tales que le permitan desenvolverse con eficiencia.

Además de los resultados teóricos y prácticos, el proyecto persigue la realización de un desarrollo experimental.

Desde la mismísima etapa de diseño los algoritmos criptográficos deben demostrar su resistencia a los ataques conocidos. Es por ello que existen instancias y funciones que deben probar su resistencia a tal o cual ataque. En particular a la enorme cantidad de nuevas y poderosas herramientas de Criptoanálisis que se han desarrollado en los últimos tiempos.

Palabras Clave:

Criptografía, Criptoanálisis, Criptosistemas de Clave Privada, Stream Ciphers. Sistemas Móviles.

CONTEXTO.

En el marco de la carrera de grado de Ingeniería en Informática y el posgrado en Criptografía y Seguridad Teleinformática que se dictan en la *Escuela Superior Técnica "Gral. Div. Manuel N. Savio" (EST)*, dependiente de la *Facultad del Ejército, Universidad de la Defensa Nacional (UNDEF)* se llevan adelante tareas de I+D+i

por parte del *Grupo de Investigación en Criptología y Seguridad Informática (GICSI)*. El mismo depende del *Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática (CriptoLab)* perteneciente al *Laboratorio Informática (InforLab)*. Y está conformado por docentes investigadores, profesionales técnicos y alumnos de dicha área.

1. INTRODUCCIÓN.

Los sistemas "fijos" de comunicaciones disponen de mayores recursos de Hardware y Software que sus pares "móviles". Estos últimos permiten la realización de comunicaciones en posiciones fijas, como también en movimiento (por ejemplo sistemas de tipo VHF¹ entre otros). Sus capacidades de almacenamiento y procesamiento (entre otras) pueden resultar fácilmente sacrificables y ser disminuidas en beneficio de una mayor movilidad, menor peso y consumo energético, entre otras limitaciones que pudieren tener.

Sin embargo, la confidencialidad de los enlaces no puede ni debe ser reducida o eliminada por los diseñadores, fabricantes o usuarios de estos equipos en persecución de una mayor libertad de movimientos. Ello implica el pago de un alto precio al atentar contra la seguridad de la comunicación.

¹ Very High Frequency: rango de frecuencias de 30 MHz a 300 MHz. Empleado por sistemas satelitales, televisión, radiodifusoras de FM, bandas aéreas y marítimas, entre otras.

La Criptografía tiene en su haber multiplicidad de algoritmos robustos y confiables. Lamentablemente no todos ellos pueden implementarse de manera eficaz y segura en Sistemas Móviles debido a la reducción de recursos con las que funcionan: un algoritmo podría demandar una cantidad de recursos (tiempo de ejecución, energía, memoria, potencia de cálculo, etc.) que el sistema sería incapaz de suministrar. Incluso asumiendo la posibilidad que pudiera hacerlo, tal vez su ejecución podría comprometer la velocidad del sistema provocando demoras y demás efectos negativos sobre la comunicación.

En esta dirección y en el ámbito de las fuerzas armadas argentinas, CriptoLab ha realizado algunas propuestas. En especial para vehículos aéreos no tripulados del Proyecto LIPAM[1] del Ejército Argentino, los cascos de realidad aumentada del Proyecto² RAIOM[2] del Centro de Investigaciones para la Defensa - CITEDEF³. También se pueden mencionar otros sistemas y vehículos militares, como el PANHARD[3] francés que el Ejército y otras fuerzas poseen y que le fue encomendado a la EST para su modernización.

Cabe aclarar que en el ámbito civil también existen dispositivos y vehículos que requieran cifrado de características similares. Es por ello que el estudio de dichos sistemas redunda en un beneficio dual: civil y militar.

El diseño e implementación de un criptosistema reducido, veloz y económico en el consumo de recursos sería indispensable para dotar de confidencialidad a datos y canales de comando y control.

A su vez investigar sobre la existencia de problemas que generan debilidades en el cifrado. Como así también someter al sistema a los ataques criptoanalíticos reconocidos como el Criptoanálisis Diferencial, Lineal, Algebraico, Cube Attack, entre otros [4-7].

² RAIOM: Realidad Aumentada para la Identificación de Objetivos Militares.

³ CITEDEF: El Instituto de Investigaciones Científicas y Técnicas para la Defensa; ex Instituto de Investigaciones Científicas y Técnicas de las Fuerzas Armadas (CITEFA)

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN.

Hemos dividido el proyecto en 4 etapas de investigación y desarrollo:

- a) Estudio y análisis de algoritmos que satisfacen los requerimientos y condiciones de entorno del proyecto.
- b) Personalización, diseño y desarrollo del algoritmo:
 - Estudio de sus vulnerabilidades y ataques conocidos.
 - Implementación y pruebas del algoritmo.
- c) Determinación de las propiedades criptológicas:
 - Estudio de las propiedades.
 - Experiencias de laboratorio.
- d) Ejecución de los tests y demás pruebas de robustez criptológica.
 - Diseño y programación de los tests.
 - Diseño e implementación de los ataques.
 - Análisis de los resultados obtenidos.
 - Redacción del informe final.
 - Puesta a punto del algoritmo a entregar.

3. RESULTADOS Y OBJETIVOS.

Se espera realizar el diseño de un esquema de cifrado y descifrado del tipo Stream Cipher (cifrado en flujo o cadena de bits) de Clave Privada, que pueda garantizar la seguridad de las comunicaciones sobre uno o más canales de comunicaciones de un Sistema Móvil.

Su fortaleza se podrá observar a través de sus propiedades matemáticas pertinentes. Además, la Secuencia Cifrante (Key Bit Stream) [8] que se obtenga, deberá satisfacer todos los requisitos aceptados por la comunidad científica que deben tener las Secuencias Seudo-Aleatorias: Test de Golomb, de NIST, Die Hard y demás, estudio de la longitud de recursión, complejidad lineal y período[9].

La realización de un desarrollo propio y nacional redundará en ahorrar recursos económicos enfrentados a los altos costos en equipos y algoritmos adquiridos en el exterior y pagaderos en moneda foránea.

4. FORMACIÓN DE RECURSOS HUMANOS.

Los docentes investigadores participantes del proyecto dictan las asignaturas *Criptografía y Seguridad Teleinformática*, *Matemática Discreta* y *Paradigmas de Programación I, II*. Desde esas cátedras se invita a los alumnos a participar en los proyectos de investigación. Es por ello que los alumnos *Dorado*, *Mariano*, *Cabrera Ezequiel*, *Leiras Facundo* y *Romero*, *Luciano* han demostrado su interés y se han sumado en calidad de colaboradores. En particular los dos últimos serán postulantes para la beca “Estímulo a las Vocaciones Científicas” (EVC) otorgadas por el Consejo Interuniversitario Nacional (CIN) por encuadrarse en las condiciones pedidas [10].

Se desea destacar que el incremento del Know-How que tendrá el grupo de investigadores a lo largo de la vida del proyecto será una importante y económica Formación de Recursos Humanos en beneficio de sus integrantes.

Atendiendo a la responsabilidad ética y social que compete a la actividad científica y tecnológica, el Grupo Integrante de este Proyecto de Investigación, ya sea durante su ejecución o por la aplicación de los resultados obtenidos, desea expresar su compromiso a no realizar cualquier actividad personal o colectiva que pudiera afectar los derechos humanos, o ser causa de un eventual daño al medio ambiente, a los animales y/o a las generaciones futuras.

5. BIBLIOGRAFÍA

[1] https://es.wikipedia.org/wiki/Lip%C3%A1n_M3 consultada el 12/3/2018.

[2] <http://www.historialprensa.mindef.gov.ar/articulos/ver/120>. consultada el 14/3/2017.

[3] Ding C.; *The differential cryptanalysis and design of natural stream ciphers*. In: Anderson R. (eds.) Fast Software Encryption. FSE 1993. Lecture Notes in Computer Science, vol. 809. Springer Berlin, Heidelberg.

[4] https://es.wikipedia.org/wiki/Panhard_AML consultada el 14/3/2018.

[5] Wu H., Preneel B. *Differential Cryptanalysis of the Stream Ciphers Py, Py6 and Pypy*. In: Naor M. (eds.) Advances in Cryptology. EUROCRYPT 2007. Lecture Notes in Computer Science, vol. 4515. Springer Berlin, Heidelberg. 2007.

[6] Muller F., Peyrin T. *Linear Cryptanalysis of the TSC Family of Stream Ciphers*. In: Roy B. (eds.) Advances in Cryptology - ASIACRYPT 2007. Lecture Notes in Computer Science, vol. 3788. Springer, Berlin, Heidelberg. 2005.

[7] Dinur I., Shamir A. *Cube Attacks on Tweakable Black Box Polynomials*. Advances in Cryptology - EUROCRYPT 2009. Lecture Notes in Computer Science, vol 5479. Springer, Berlin, Heidelberg. 2009

[8] Pasalic, E.; *On Guess and Determine Cryptanalysis of LFSR-Based Stream Ciphers*; IEEE Transactions on Information Theory. Vol. 55 Ed.7°, 2009.

[9] Biryukov A., Shamir A. (2000) Cryptanalytic Time/Memory/Data Tradeoffs for Stream Ciphers. In: Okamoto T. (eds) Advances in Cryptology — ASIACRYPT 2000. ASIACRYPT 2000. Lecture Notes in Computer Science, vol 1976. Springer, Berlin, Heidelberg.

[10] <http://evc.cin.edu.ar/informacion> consultada el 23/2/2018.

Toolkit para Criptoanálisis

Castro Lechtaler, Antonio^{1,2}; Cipriano, Marcelo^{1,3}; García, Edith¹,
Liporace, Julio¹; Maiorano, Ariel¹; Malvacio, Eduardo¹; Tapia, Néstor¹;

¹Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática.
Escuela Superior Técnica, Facultad del Ejército. Universidad de la Defensa Nacional - UNDEF

²CISTIC/FCE - Universidad de Buenos Aires.

³Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes UNQ.

acastro@est.iue.edu.ar , marcelocipriano@est.iue.edu.ar,
{edithgarcia; jcliporace; maiorano; edumalvacio; tapianestor87}@gmail.com

RESUMEN

El proyecto se extiende desde el estudio hasta la implementación de técnicas o métodos criptoanalíticos para ser aplicados a determinados generadores de secuencias pseudoaleatorias tipo *Stream Ciphers*¹, en particular a aquellos algoritmos que involucran *LFSR*'s², *NLFSR*'s³, *CCG*'s⁴ y *CA*'s⁵.

En particular los métodos a estudiar serán los correspondientes al *Criptoanálisis Diferencial* [1-2], *Criptoanálisis Lineal* [3], *Criptoanálisis Algebraico*, *Guess-and-Determine* [4] y una de las últimas y poderosas técnica de criptoanálisis publicada hace relativamente poco, denominada *Cube Attack*⁶ [5].

El desarrollar un conjunto de herramientas de criptoanálisis permitirá la realización de análisis de algoritmos de cifrado, generadores de secuencias pseudoaleatorias, primitivas criptológicas, protocolos de seguridad de la información, y claves secretas de distintos criptosistemas, entre otros.

Palabras Clave

Criptología, Criptoanálisis. Stream Ciphers.

CONTEXTO

En el marco de la carrera de grado de Ingeniería en Informática y el posgrado en Criptografía y Seguridad Teleinformática que se dictan en la *Escuela Superior Técnica "Gral. Div. Manuel N. Savio" (EST)*, dependiente de la *Facultad del Ejército, Universidad de la Defensa Nacional (UNDEF)* se llevan adelante tareas de I+D+i por parte del *Grupo de Investigación en Criptología y Seguridad Informática (GICSI)*.

GICSI depende del *Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática (CriptoLab)* perteneciente al *Laboratorio Informática (InforLab)*. Y está conformado por docentes investigadores, profesionales técnicos y alumnos de dicha área.

1. INTRODUCCIÓN

Ya es claro que el uso masivo de las comunicaciones electrónicas ha cambiado nuestra sociedad. Desde computadoras y teléfonos inteligentes a dispositivos *smart* como televisores, electrodomésticos y otros sistemas que anexaron la comunicación a sus funciones. Es por ello que la transmisión y almacenamiento de información requiere tomar una serie de medidas de protección manteniendo su

¹ Stream cipher: generadores pseudoaleatorios conocidos también como generadores en flujo o cadena.

² Linear Feedback Shift Registers: registros de desplazamiento realimentados linealmente.

³ Non Linear Feedback Shift Registers: registros de desplazamiento realimentados no linealmente.

⁴ Clock Controlled Generators: generadores controlados por reloj.

⁵ Cellular Automata: autómata celular.

⁶ Presentado en el congreso EuroCrypt del año 2009 por sus autores: Itai Dinur y Adi Shamir.

confidencialidad, autenticidad e integridad. La falta total o parcial de medidas de seguridad se convierte en una amenaza latente.

Al momento de realizar el diseño de un criptosistema se deben tener en cuenta, obviamente, todos los ataques que éste puede sufrir. Cada aspecto en su diseño responde, entre otros criterios, a un ataque criptoanalítico, demostrando así su resistencia a él. Esta filosofía de diseño ha llevado a los modernos criptológicos a demostrar que pueden sortear diferentes amenazas, cuidadosa y eficientemente desarrolladas para atacarlos. Ya que cada algoritmo, cada primitiva, cada protocolo debe ser atacado mediante una técnica adecuada a su estructura.

Es por ello que existen diferentes y poderosas herramientas de criptoanálisis. Y todo parece indicar que se incorporarán cada vez más de ellas, en un futuro no muy lejano. De la misma manera se requerirán nuevos y probados sistemas criptográficos que resistan los ataques que van apareciendo. Y así seguirá.

Últimamente el diseño de algoritmos seguros ha tenido un gran avance. De un tiempo a esta parte, entre otros:

- el llamado en 1997 del *NIST*⁷ para escoger un nuevo algoritmo como estándar de cifrado llamado *AES* [6].
- El concurso europeo *e-Stream* en 2004, organizado por el *E-CRYPT* [7] del cual superaron todas las pruebas y ataques, 7 algoritmos.
- el llamado a concurso del *NIST* para escoger un nuevo algoritmo como estándar *SHA-3*⁸, finalizado en 2012. Aunque aún no se da de

baja al *SHA-2* por no demostrar, hasta el momento, debilidades.

- el concurso aún en proceso *CAESAR*⁹ el cual se espera que pronto se den a conocer a los finalistas y luego al ganador o un portfolio con los algoritmos elegidos[8].

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Se ha dado en planificar este proyecto de investigación siguiendo 6 etapas:

1. Mediante el estudio de bibliografía actualizada y la asistencia a Cursos, Congresos y Workshops específicos, se profundizará en el estado del arte del Criptoanálisis y los nuevos ataques que se han desarrollado.
2. Estudio, análisis y selección de los generadores de secuencias cifrantes.
3. Relevamiento de los métodos criptoanalíticos que se analizarán.
4. Estudio de técnicas criptográficas para la determinación del o los métodos de ataque adecuados a la estructura del algoritmo estudiado.
5. Implementación de los métodos de criptoanálisis.
6. Análisis de los resultados obtenidos.

3. RESULTADOS OBTENIDOS / ESPERADOS

Se realizará el estudio, análisis y desarrollo de técnicas y/o herramientas que posibiliten la realización del diseño de aplicaciones criptográficas, su evaluación, búsqueda de

⁷ Institución de Estados Unidos, llamada Instituto de Normas y Estandarización (National Institute of Standards and Technology) por sus siglas en inglés.

⁸ Secure Hash Algorithm por sus siglas en inglés.

⁹ CAESAR: Competition for Authenticated Encryption: Security, Applicability, and Robustness.

vulnerabilidades y de ser posible, lograr su ataque.

Los alcances del criptoanálisis podrán ser:

- Obtención de la/s clave/s del cifrado.
- Hallar patrones estadísticos en la salida del sistema estudiado.
- Desarrollar nuevas técnicas criptoanalíticas de acuerdo a las propiedades del sistema estudiado.
- Analizar el algoritmo de generación de la/s clave/s y estudiar su vulnerabilidad.

Para ello se perseguirán los objetivos particulares:

- Estudio y análisis de técnicas criptoanalíticas.
- Diseño y desarrollo de herramientas de evaluación, ataque o quiebre de aplicaciones criptográficas.
- Pruebas y testeo de las herramientas desarrolladas sobre algoritmos específicos.

4. FORMACIÓN DE RECURSOS HUMANOS

Los docentes investigadores participantes del proyecto dictan las asignaturas *Criptografía y Seguridad Teleinformática, Matemática Discreta y Paradigmas de Programación I, II*. Desde esas cátedras se invita a los alumnos a participar en los proyectos de investigación. Es por ello que los alumnos *Dorado, Mariano, Cabrera Ezequiel, Leiras Facundo y Romero, Luciano* han demostrado su interés y se han sumado en calidad de colaboradores. En particular los dos últimos serán postulantes para la beca "Estímulo a las Vocaciones Científicas" (EVC) otorgadas por el Consejo Interuniversitario Nacional (CIN) por encuadrarse en las condiciones pedidas [9].

Se desea destacar que el incremento del Know-How que tendrá el grupo de investigadores a lo largo de la vida del proyecto será una importante y económica Formación de Recursos Humanos en beneficio de sus integrantes.

Atendiendo a la responsabilidad ética y social que compete a la actividad científica y tecnológica, el Grupo Integrante de este Proyecto de Investigación, ya sea durante su ejecución o por la aplicación de los resultados obtenidos, desea expresar su compromiso a no realizar cualquier actividad personal o colectiva que pudiera afectar los derechos humanos, o ser causa de un eventual daño al medio ambiente, a los animales y/o a las generaciones futuras.

5. BIBLIOGRAFÍA

- [1] Ding C.; *The differential cryptanalysis and design of natural stream ciphers*. In: Anderson R. (eds.) *Fast Software Encryption. FSE 1993. Lecture Notes in Computer Science*, vol. 809. Springer Berlin, Heidelberg.
- [2] Wu H., Preneel B. *Differential Cryptanalysis of the Stream Ciphers Py, Py6 and Pypy*. Naor M. (eds.) *Advances in Cryptology. EUROCRYPT 2007. Lecture Notes in Computer Science*, vol. 4515. Springer Berlin, Heidelberg. 2007.
- [3] Muller F., Peyrin T. *Linear Cryptanalysis of the TSC Family of Stream Ciphers*. Roy B. (eds.) *Advances in Cryptology - ASIACRYPT 2007. Lecture Notes in Computer Science*, vol. 3788. Springer, Berlin, Heidelberg. 2005.
- [4] Pasalic, E.; *On Guess and Determine Cryptanalysis of LFSR-Based Stream Ciphers*; *IEEE Transactions on Information Theory*. Vol. 55 Ed.7°, 2009.
- [5] Dinur I., Shamir A. *Cube Attacks on Tweakable Black Box Polynomials*. *Advances in Cryptology - EUROCRYPT 2009. Lecture Notes in Computer Sci-*

ence, vol 5479. Springer, Berlin, Heidelberg. 2009.

[6] Daemen, J.; Rijmen, V.; *The Design of Rijndael: AES - The Advanced Encryption Standard*. Springer. New York. 2002.

[7] <http://www.ecrypt.eu.org/stream/>
Consultada el 10-3-18.

[8] <https://competitions.cr.yo.to/caesar.html>. Consultada el 10-3-18.

[9] <http://evc.cin.edu.ar/informacion>
consultada el 23/2/2018.

Análisis de Vulnerabilidades de Sistemas Web en desarrollo y en producción

Juan C. Cuevas, Roberto M. Muñoz, M. Alejandra Di Gionantonio,
Iris N. Gastañaga, Fabián A. Gibellini, Germán Parisi, Diego Barrionuevo, Milagros Zea
Cárdenas.

Laboratorio de Sistemas / Dpto. de Ingeniería en Sistemas / Universidad Tecnológica
Nacional / Facultad Regional Córdoba. Cruz Roja S/N, 5016

juancarloscue@hotmail.com, robertomunioz@gmail.com, ing.alejandradg@gmail.com,
irisg@ciec.com.ar, fgibellini@bbs.frc.utn.edu.ar, germannparisi@gmail.com,
santosdiegob@gmail.com, milyzc@gmail.com

Resumen

La realización de pruebas de penetración permiten detectar las vulnerabilidades de los sistemas de información, en el caso particular del proyecto de investigación y desarrollo que se lleva a cabo en UTN FRC, se han realizado acciones de esta metodología en sistemas web que desarrollan los estudiantes del último año de la carrera de Ingeniería en Sistemas de Información y con acciones similares que se realizaron en Sistemas web en producción de empresas privadas nacionales e internacionales.

Se presentará el abordaje teórico, la metodología utilizada y las técnicas con las que se llevan a cabo diversas pruebas manuales, se muestran los resultados obtenidos de aplicar la metodología para luego sentar las bases en búsqueda de la repetición de dichas pruebas ante un mismo sistema objetivo generando su automatización.

Palabras clave: Seguridad web; Auditorías; Ethical Hacking; Vulnerabilidad; Pentesting.

Contexto

Este proyecto de investigación se lleva a cabo en el ámbito del Laboratorio de Sistemas (LabSis) de la Universidad Tecnológica

Nacional – Facultad Regional Córdoba (UTN - FRC).

Se enmarca dentro de las líneas Seguridad Informática y forma parte del proyecto de investigación Sistema Integrado de Soporte para Análisis de Vulnerabilidades en Sistemas Web (S.I.S.A.V.S.W.). Código: EIUTNCO0004084. Acreditado por la Secretaría de Ciencia, Tecnología y Postgrado, y financiado por la UTN – FRC.

Introducción

La seguridad de la información describe actividades relativas a la protección de la información y los activos de la infraestructura de la información contra riesgos de pérdida, uso inadecuado, revelación o daño.

Los riesgos de estos activos pueden ser calculados mediante el análisis de las siguientes cuestiones:

- Amenazas a sus activos: Eventos no deseados que pueden causar pérdida, daño o uso inadecuado de los activos en forma deliberada o accidental.
- Vulnerabilidades: Se refiere a cuán susceptibles son sus activos a ataques.
- Impacto: La magnitud de la pérdida potencial o la seriedad del evento.

Existen disponibles normas, modelos y estándares para asistir a las organizaciones en la implementación de programas y controles apropiados para mitigar estos riesgos, como por

ejemplo, las normas ISO como por ejemplo la ISO 27.001 [1], los modelos ITIL [2], COBIT [3] y estándares del NIST [4].

En lo referido a las vulnerabilidades que es de interés para las pruebas de penetración diferentes autores abordan la temática.

Xie et al postulan que muchas de las vulnerabilidades de seguridad en las aplicaciones actuales son introducidas por los desarrolladores de software al escribir código inseguro los cuales se pueden deber a la falta de comprensión de programación segura [6].

Bates et al postulan que en los últimos años, los fabricantes de navegadores e investigadores han tratado de desarrollar filtros del lado del cliente para mitigar estos ataques. Pero algunos de estos filtros podrían introducir vulnerabilidades en sitios que antes estaban libres de errores. [7]

Tripp et al postulan que los autores de publicaciones recientes sugieren que las aplicaciones web son altamente vulnerables a ataques de seguridad. En referencia a lo cual citan un reporte reciente de la WASC que proveen estadísticas de seguridad sobre 12186 sitios web en producción, listando un total de 97554 vulnerabilidades detectadas en estos sitios web. Más severamente aún es que cerca del 49% de los sitios analizados se encontró que contenían vulnerabilidades de alto riesgo [8].

Según Shahriar las aplicaciones web contienen vulnerabilidades, las cuales pueden conducir a serias brechas de seguridad tal como el robo de información confidencial [9].

Grossman en un reporte indica que las aplicaciones web de diferentes dominios de uso más frecuente (por ejemplo: banca, salud, TI, educación, redes sociales) son más propensas a ser vulneradas. [10]

Ha habido una gran puja por incluir la seguridad durante el ciclo de desarrollo de software y formalizar la especificación de requerimientos de una manera estándar. Además de un gran incremento en la cantidad de organizaciones dedicados a la seguridad de aplicaciones como la Open Web Application Security Project (OWASP) [11]. Sin embargo,

todavía hay aplicaciones web descaradamente vulnerables, principalmente porque los programadores están más concentrados en la funcionalidad que en la seguridad [12].

En referencia a los ataques, según el último top 10 de las OWASP realizado cada 3 años, esta organización lista los ataques más usados para explotar vulnerabilidades. Estos son: Inyección (SQL, OS y LDAP), pérdida de autenticación y gestión de sesiones, cross site scripting (XSS), referencia directa e insegura a objetos, configuración de seguridad Incorrecta, exposición de datos sensibles, inexistente control de acceso a funcionalidades, falsificación de peticiones en sitios cruzados (CSRF), uso de componentes con vulnerabilidades conocidas, reenvíos y redirecciones no válidas. [5]

Respecto a las metodologías que son pasibles de ser utilizadas, existen mejores prácticas sobre las cuales basar la realización de este tipo de evaluación, aunque en general, cada profesional puede incorporar sus variantes. Algunos ejemplos pueden ser: el documento del NIST [4], el documento del Open Source Security Testing Methodology Manual (OSSTMM) [13], el marco de trabajo denominado Information System Security Assessment Framework (ISAAF) [14] y el Open Web Application Security Project (OWASP) [11].

Además existen herramientas que facilitan a los pentester realizar ataques de Injection, particularmente, existen varias relacionadas al ataque SQL Injection como por ejemplo: SQLmap, Havij y V1p3R. Estas herramientas se caracterizan principalmente por contener vectores de ataques más usados permitiendo automatizar un ataque [14] [15].

Cross-Site Scripting (XSS) aborda otro tipo de ataque común en aplicaciones web que consisten en aprovechar las características de los lenguajes que se ejecutan en el navegador, tales como el lenguaje JavaScript. Yusof y Pathan referencian tres tipos de ataques XSS: persistente, no persistente y basado en el DOM [16]. Hay varias maneras de prevenir estos

ataques XSS: mediante el “sanitizado” de todas las entradas de información al sistema [16] o un método que haga inútil la cookie que pueda robar el atacante [17].

En los siguientes puntos se define la metodología que se lleva a cabo para realizar las pruebas y obtener información para la generación de las bases que permitirán el desarrollo de un sistema automatizado de pruebas de penetración para los sistemas objetivos.

Metodología y técnicas

Entre las actividades desarrolladas en este proyecto orientadas a abordar esta temática se procedió a realizar una pormenorizada identificación y clasificación de técnicas y herramientas. Con respecto a estas últimas se realizó una búsqueda y selección de las mismas en el mercado, bibliografía, publicaciones científicas y en la web durante el 2016 - 2017. Inicialmente lo único que tenían que cumplir las herramientas era que fueran de software libre. Los criterios de selección fueron empíricos, ya que los integrantes del equipo de pentesters analizaron sus salidas en pruebas de seguridad realizadas a empresas privadas y a desarrollos de proyectos de la materia Proyecto Final en la carrera de Ingeniería en Sistemas de Información de la UTN FRC, verificaban si cada herramienta probada era lo suficientemente confiable y sus salidas eran comprensibles. Si cumplían con estos criterios, la herramienta era utilizada por los pentesters en próximas pruebas de seguridad.

Se realizaron trabajos de pentest a un proyecto de la materia de proyecto del último año de la carrera de Ingeniería en Sistemas de Información de la UTN FRC, con metodología de ethical hacking y dos pentest a una empresa privada nacional y otra internacional. Llevando a cabo las mismas acciones en todas:

- **Visibilidad** de tipo caja negra: El tester de seguridad no cuenta con ninguna

información del objetivo. El tester de seguridad no tiene accesos, la única información sobre vulnerabilidades serán las que se puedan determinar mediante técnicas de prueba y “adivinación”.

- **Posicionamiento** Externo: El posicionamiento externo brinda al tester de seguridad la posición de atacantes externos a la organización de donde suelen provenir la mayor cantidad de ataques, en la mayoría de las organizaciones el nivel de madurez de la protección perimetral es alto. El tester de seguridad realiza las pruebas desde fuera de la organización. Busca probar el perímetro expuesto a internet.
- **Etapas del análisis de seguridad:**
 - Reconocimiento Pasivo
 - Reconocimiento activo superficial.
 - Reconocimiento activo en profundidad.
 - Análisis de Vulnerabilidades.

De lo expuesto anteriormente se obtuvo como resultados los siguientes valores expresados en una tabla.

La tabla 1 identifica las categorías, basadas en OWASP y de acuerdo a las etapas del ciclo de desarrollo donde se han identificado las vulnerabilidades.

En la tabla no se encuentran todas categorías definidas por OWASP debido a que muchas de estas no han sido detectadas como vulnerabilidades en los sistemas objetivos.

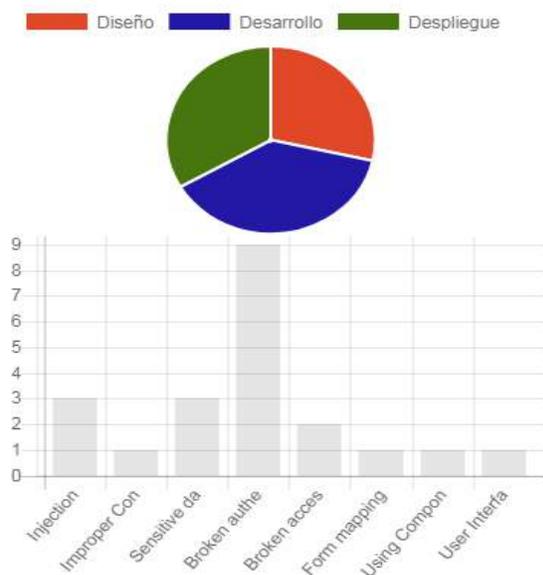
En el caso de la categoría Form Mapping, no se encuentra identificada en las OWASP, pero se pudo detectar vulnerabilidades en los frameworks web que se analizaron.

Tabla 1

Categoría	Diseño	Desarrollo	Despliegue	Total
Injection	0	3	0	3
Improper Control of Interaction Frequency	0	1	0	1
Sensitive data exposure	1	1	1	3
Broken authentication	5	0	4	9
Broken access control	0	2	0	2
Form mapping	0	1	0	1
Using Components with Known Vulnerabilities	0	0	1	1
User Interface Misrepresentation of Critical Information	0	0	1	1
Total	6	8	7	21

Se puede observar (Gráfico 1) que los mecanismos de desarrollo de autenticación son los que más fallan y se nota que hay un pico alto de bugs de inyecciones y esto concuerda con el top de 2017 de OWASP

Gráfico 1



El resto de las vulnerabilidades detectadas a pesar de que en números es menor, no deja de ser un llamado de atención para los líderes de estos sistemas.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto se inscribe dentro de esta línea de investigación de seguridad de la

información, más específicamente pruebas de penetración en sistemas web en producción, de forma que permitan identificar las vulnerabilidades existentes dentro de un ambiente controlado.

Resultados y Objetivos

Como resultado del análisis surgen algunas reflexiones primigenias: Los Sistemas web objetivos de estudio poseen vulnerabilidades bien conocidas, ya sea en los desarrollos de los estudiantes como en el ámbito profesional privado en empresas nacionales e internacionales. Así mismo, es necesario tener en cuenta que las acciones que se deben llevar a cabo para realizar estas pruebas y poder obtener datos fehacientes dependen de un acuerdo de confiabilidad y confidencialidad de los propietarios del sistema con el equipo que lleva a cabo la metodología de ethical hacking por lo que se dificulta poder contar con gran cantidad de muestras para poder llegar a conclusiones definitivas, pero las que aquí se pudieron obtener constituyen, el comienzo de un camino que permita definir las bases de una sistema automatizado de pruebas de penetración. Otro aspecto que se suma también como dificultad son los tiempos que demanda llevar estas acciones sobre los sistemas web.

En la práctica, a partir de experiencias de los integrantes del equipo se identificaron, por un lado, la necesidad de trabajar fuertemente en incluir los requerimientos de seguridad en la planificación de nuevos proyectos de desarrollo de productos software, como así también la necesidad de gestionar la ejecución de múltiples pruebas de penetración en el contexto de la seguridad de la información de sistemas web en producción, basándose en metodologías abiertas, para identificar y analizar sus vulnerabilidades. Para lo que es necesario:

- Permitir a los pentesters identificar vulnerabilidades y automatizar el proceso de identificación.

- Lograr que el sistema emita un diagnóstico respecto a las vulnerabilidades del sistema web analizado.
- Crear una base de datos que facilite al pentester vincular metodologías, técnicas y herramientas para abordar la evaluación de vulnerabilidades de un sistema web determinado.
- Generar un sistema que contribuya y facilite el desarrollo de las actividades del pentester.
- Obtener un producto (sistema) que sea simple de utilizar por los profesionales de la seguridad.

Formación de Recursos Humanos

En el equipo trabajarán estudiantes avanzados de la carrera de Ingeniería en Sistemas de Información, los cuales actualmente se desempeñan en el Laboratorio de Sistemas (LabSis), con la finalidad de que inicien su formación en investigación científica y tecnológica profundizando sus conocimientos en temas significativos en la seguridad de la información. Los estudiantes podrán realizar la Práctica Supervisada.

Se llevan a cabo competencias de casos de test de penetración ético, para fomentar en los estudiantes el pensamiento en el desarrollo seguro, e inculcar conceptos de ética ante estas acciones.

Se realizaron capacitaciones a estudiantes sobre cómo resolver vulnerabilidades bien conocidas, a partir de realizar pruebas de penetración ético.

Referencias

- [1] ISO/IEC 27001. "Tecnología de la información". Técnicas de la seguridad. Sistemas de gestión de la seguridad de la información (SGSI). Requisitos. ISO Ginebra, Suiza 2013.
- [2] ITIL. Information Technology Infrastructure Library. Vs. 3 - 2001.
- [3] COBIT Control Objectives for Information and Technology. Vs. 5 - 2012.
- [4] Technical Guide to Information Security Testing and Assessment. SP 800-115. NIST

National Institute of Standards and Technology. 2008.

[5] "OWASP top 10 2013 Project". Open Web Application Security Project. https://www.owasp.org/index.php/Top_10_2013-Top_10. 2013.

[6] XIE, J.; Chu B.; Lipfort, H. R.; Melton, J. T.: "ASIDE: IDE Support for Web Application Security". ACSAC '11, Orlando, Florida USA. Dec. 5-9/2011

[7] Bates,D; Barth, A.; Jackson, C.: "Regular Expressions Considered Harmful in Client-Side XSS Filters", Raleigh, NC, USA. April 26-30/2010.

[8] Tripp, O.; Weisman, O. y Guy, L.: "Finding Your Way in the Testing Jungle". ISSTA '13, July 15-20/2013, Lugano, Switzerland.

[9] Shahriar, H.: "Security Vulnerabilities and Mitigation Techniques of Web Applications". SIN'13, November 26-28/2013, Aksaray, Turkey.

[10] Grossman, J.: "How does your website security stack up against peers?" White Hat Report, Summer. 2012.

[11] Guía OWASP (Open Web Application Security Project). Vs. 3.

[12] Josh Pauli: "The Basics of Web Hacking: Tools and Techniques to Attack the Web". Ed. ELSEVIER. 2013. 25 Wymnan Street, Waltham, MA 02451, USA.

[13] "Open Source Security Testing Methodology Manual (OSSTMM)". Institute for Security and Open Methodologies (ISECOM). Diciembre 2010. Cataluña. España.

[14] Ciampa, A.; Visaggio, C. A.; y Di Penta, M. "A heuristic-based approach for detecting SQL-injection vulnerabilities in Web applications". SESS. Cape Town, South Africa. Mayo 2/2010.

[15] Nagpal, B; Chauhan, N.; Singh, N.; Panesar, A.: "Tool Based Implementation of SQL Injection for Penetration Testing". International Conference on Computing, Communication and Automation (ICCCA2015). 2015

[16] Yusof, I; Pathan, A. S.; "Preventing Persistent Cross-Site Scripting (XSS) Attack By Applying Pattern Filtering Approach". IEEE. 2014.

[17] Takahashi H.; Yasunaga K.; Mambo M.; Kim K.; Youl Youm H.: "Preventing Abuse of Cookies Stolen by XSS". Eighth Asia Joint Conference on Information Security. IEEE 2013.

Vulnerabilidades en HTTP/2

Mauro F. Fornaroli ⁽¹⁾, Mónica D. Tugnarelli ⁽¹⁾, Sonia R. Santana ⁽¹⁾, Javier Díaz ⁽²⁾

⁽¹⁾Facultad de Ciencias de la Administración – Universidad Nacional de Entre Ríos

⁽²⁾Facultad de Informática – Universidad Nacional de La Plata

e-mail: maufor, montug [@fcad.uner.edu.ar]

Resumen

El PID 7052 tiene por objetivo obtener conclusiones generales y comparativas acerca de dos metodologías de recolección de datos digitales, una a priori y otra a posteriori de un evento de seguridad analizando particularmente su performance en entornos de servidores web. Como parte de las actividades previstas en este proyecto se realiza un análisis exploratorio del protocolo HTTP, comparándose las versiones HTTP 1.0 y 1.1 y la nueva HTTP/2 e identificando puntos de control sobre los cuales se realiza la captura de tráfico y recolección de datos. Se consideran aspectos de seguridad, concentrándose en el estudio de vulnerabilidades conocidas de HTTP/2. Además, se define y configura un entorno de testing, utilizando herramientas Open Source, para simular un ataque de Denegación de Servicio Distribuido (DDoS) a los fines de observar el comportamiento del servidor y el volumen de datos resguardados en logs y para analizar la performance de las metodologías mencionadas en instancias de recuperación del servicio y resguardo de la evidencia digital.

Palabras clave: HTTP, seguridad, vulnerabilidades, DDoS

Contexto

El Proyecto de Investigación y Desarrollo PID-UNER 7052 para Director Novel con Asesor, denominado “Análisis de Metodologías de Recolección de datos digitales” se encuadra en una de las líneas de investigación establecidas como prioritarias para su fomento, "Arquitectura, Sistemas Operativos y Redes", de la carrera Licenciatura en Sistemas de la Facultad de Ciencias de la Administración. Se adecua, además, a las prioridades de la Universidad Nacional de Entre Ríos por ser un proyecto aplicado a la investigación sobre Tecnologías de la Información y la Comunicación [1].

Introducción

HyperText Transfer Protocol (HTTP) es un protocolo de nivel de aplicación con características definidas para utilizarse en sistemas de información distribuidos, colaborativos e hipermediales. Se han introducido mejoras en sus diferentes versiones sobre todo tendientes a optimizar la performance del mismo, reducir el consumo de recursos y la latencia y para resolver algunos inconvenientes que tiene la comunicación a través TCP [2][3][4][5].

La última versión HTTP/2 [6], presenta un protocolo binario que incorpora

multiplexación y el uso obligatorio de TLS (Transport Layer Security) conservando la misma semántica y la compatibilidad con las versiones 1.0 y 1.1. El protocolo se implementa si el cliente y el servidor tienen soporte y en el caso de que alguno de los dos no lo tengan, en la negociación de protocolo, se acuerda usar las versiones anteriores. Actualmente la mayoría de los browsers y entornos de servidor cuentan con implementaciones oficiales para esta nueva versión.

Entre las características más relevantes del protocolo HTTP/2 pueden mencionarse [6][7][8]:

- Formato de mensajes basados en tramas;
- Reducción de la latencia al permitir una multiplexación completa de streams de solicitudes y respuestas;
- Disminución de la sobrecarga de protocolo mediante la compresión eficiente y codificación binaria HPACK (Header Compression for HTTP/2);
- Incorporación de soporte para la prioridad de solicitudes y capacidades “Server Push”;
- Definición de nuevos mecanismos de control de flujo, manejo de errores y actualizaciones.

Dado que HTTP es considerado un protocolo no seguro, en el desarrollo de HTTP/2 se identificaron y gestionaron los riesgos de seguridad del nuevo protocolo, tanto en base a decisiones de diseño como a guías de implementación. Sin embargo, algunas implementaciones de servidores HTTP/2 pueden no seguir estrictamente estas guías, tornándose vulnerables a

distintos tipos de ataques de Denegación de Servicio (DoS) [8].

Algunas vulnerabilidades conocidas son:

- Stream Reuse,
- Slow Read Attack,
- Dependency Cycle DoS,
- HPACK Bomb.

Resulta de interés entender como este tipo de ataques puede afectar el comportamiento del servidor y para analizar la performance de las metodologías de recolección de datos digitales en instancias de recuperación del servicio y resguardo de la evidencia digital.

Líneas de Investigación, Desarrollo e Innovación

Con este proyecto de investigación se espera conformar una base de conocimiento sobre la forensia informática en relación a metodologías de recolección de datos digitales. Como parte de las actividades se desarrolla un estudio exploratorio de las principales características de HTTP como aporte para la identificación de puntos de control de este protocolo y su comportamiento ante incidentes de seguridad.

Resultados y Objetivos

El objetivo principal del PID 7052 es obtener conclusiones generales y comparativas acerca de las metodologías de recolección de datos a priori de un evento de seguridad y de recolección de datos a posteriori de un evento de seguridad,

analizando particularmente su performance en entornos de servidores web.

Con el desarrollo de las actividades previstas en el proyecto, se obtuvieron los siguientes resultados:

Identificación y descripción de puntos de control en protocolos HTTP. Se analizaron y compararon las versiones del protocolo HTTP 1.0 y 1.1 y la nueva HTTP/2, confeccionándose una tabla [9] que destaca sus principales prestaciones y diferencias entre ellas (Tabla 1). Luego, se identificaron los puntos de control para el protocolo HTTP, que en este proyecto se implementó en un servidor web Apache, sobre los cuales se realizó la captura de tráfico y recolección de datos. Se controlaron los siguientes puertos y archivos:

- puertos 80 y 443: que registra el tráfico entrante y saliente del protocolo HTTP y asociados como SSL - *Secure Sockets Layer*.
- archivos log del sistema operativo Ubuntu Linux (*/var/log/*): *messages.log*, *auth.log*, *secure*, *tmp/wtmp*
- Archivos log de Apache: *error.log* y *access.log*.
- archivos de configuración del servidor Apache

Versión	HTTP/1.0 (1996)	HTTP/1.1 (2000)	HTTP/2 (2015)
RFC	RFC 1945	RFC 2616	RFC 7540
Manejo de requerimientos	Un requerimiento entregado por vez sobre una conexión.	Mecanismo HTTP Keep Alive: varios requerimientos pueden utilizar múltiples conexiones con el servidor para reducir la latencia.	Múltiples mensajes de requerimiento/ respuesta sobre una misma conexión. Permite asignar prioridades a los requerimientos.
Header	Formato texto	Formato texto	Formato binario. - Compresión de header (Algoritmo HPACK).
Multiplexación	No permite conexiones simultáneas	No permite conexiones simultáneas	Permite múltiples solicitudes y respuestas en paralelo usando la misma conexión TCP, enviando cada requerimiento en un stream diferente.
Servidor	Descarga de recursos a solicitud del cliente (primero HTML, luego CSS, JS, imágenes, enlaces)	Descarga de recursos a solicitud del cliente (primero HTML, luego CSS, JS, imágenes, enlaces)	Tecnología server push: Permite cargar los archivos (CSS, JS, imágenes) desde el servidor al cliente sin que éste lo pida.

Tabla 1. Principales diferencias entre versiones HTTP

Definición y Configuración de entornos de testing. Se configuró un entorno de trabajo en una red LAN Ethernet conformada por un servidor, con Sistema Operativo Ubuntu Server [10] versión 15.10 y un Servidor Web Apache Server [11] versión 2.4.12. Se dispuso de una notebook dedicada con el toolkit Kali Linux versión 64bit 2017.1 [12]. Como marco de trabajo y guía general de buenas prácticas para las pruebas de laboratorio se seleccionaron las especificaciones RFC 3227 [13], ISO/IEC 27037:2012[14], OSSTMM (Open Source Security Testing Methodology Manual) [15]

Simulación de ataque de Denegación de Servicios (DoS). Empleando herramientas de hacking se realizó la simulación de un ataque DDoS, con metodología de inundación, a los fines de observar el

comportamiento del servidor, el volumen de datos resguardados en logs y para analizar la performance de las metodologías de recolección de datos digitales en instancias de recuperación del servicio y resguardo de la evidencia. Para darle la característica de distribuido, se emplearon varias máquinas virtuales actuando en paralelo. Los ataques se realizaron tanto desde el interior como del exterior de la red local.

Como avances y a modo de conclusiones parciales de estas etapas, se puede establecer que:

1. HTTP/2 permite establecer puntos de control para analizar su comportamiento ante la recolección de datos frente a un incidente de seguridad.
2. Las herramientas de forensia informática open source, como las usadas en este PID, brindan el soporte necesario y eficiente tanto para implementar un seguimiento preventivo que incluya recolección de datos como para un análisis forense luego de un incidente de seguridad.
3. Los puntos de control del protocolo HTTP, representados en su mayor parte por archivos log, muestran un crecimiento exponencial al momento en que se realizó la simulación del ataque DDoS, por lo que requiere pensar en una infraestructura adecuada.
4. Las vulnerabilidades presentes en HTTP/2 facilitan nuevas variantes de DoS. Por ejemplo, mientras que el atacante en HTTP 1.x tenía que abrir

tantas conexiones TCP como el servidor de la víctima, en HTTP / 2 los ataques se vuelven más simples, ya que se puede utilizar la multiplexación de flujos de datos.

Formación de Recursos Humanos

Este proyecto prevé la formación e iniciación en actividades de investigación de tres docentes de la carrera Licenciatura en Sistemas, la formación de dirección en proyectos de un docente y la realización de una tesis de maestría correspondiente a la Maestría en Redes de Datos de la Facultad de Informática de la UNLP.

Referencias

- [1]. Tugnarelli, M., Fornaroli, M., Santana, S., Jacobo, E., Díaz, J.: Análisis de Metodologías de Recolección de Datos Digitales. Libro de Actas Workshop de Investigadores en Ciencias de la Computación 2017, pp. 1000-1004. ISBN 978-987-42-5143-5.
- [2]. RFC 1945 Hypertext Transfer Protocol-HTTP/1.0
<http://tools.ietf.org/html/rfc1945>
- [3]. RFC 2616 Hypertext Transfer Protocol-HTTP/1.1
<http://tools.ietf.org/html/rfc2616>
- [4]. James F Kurose, Keith W. Ross. Redes de Computadoras. Un Enfoque descendente basado en Internet, 2da. Edición, Pearson Educación, 2004

- [5]. Balachander Krishnamurthy, Jeffrey C. Mogul y David M. Kristol. Key Differences between HTTP/1.0 and HTTP/1.1.
- [6]. RFC 7540 Hypertext Transfer Protocol Version 2 (HTTP/2)
<https://tools.ietf.org/html/rfc7540>
- [7]. Daniel Stenberg. HTTP2 Explained. ACM SIGCOMM Computer Communication Review, Volume 44, Number 3, July 2014.
- [8]. Imperva. Hacker Intelligence Initiative. HTTP/2: In-depth analysis of the top four flaws of the next generation web protocol. 2016
- [9]. Mónica D. Tugnarelli, Mauro F. Fornaroli, Sonia R. Santana, Eduardo Jacobo, Javier Díaz: Análisis de metodologías de recolección de datos digitales en servidores web. Libro de Actas. XXIII Congreso Argentino de Ciencias de la Computación CACIC 2017. VI Workshop de Seguridad Informática, pp. 1230-1238. ISBN 978-950-34-1539-9.
- [10]. Ubuntu Server
<https://www.ubuntu.com/server>
- [11]. Apache Server
<https://httpd.apache.org/>
- [12]. KALI Linux. www.kali.org
- [13]. RFC 3227 Guidelines for Evidence Collection and Archiving.
<https://www.ietf.org/rfc/rfc3227.txt>
- [14]. Guidelines for identification, collection, acquisition and preservation of digital evidence” ISO/IEC 27037:2012
- [15]. Open Source Security Testing Methodology Manual (OSSTMM)
<http://www.isecom.org/mirror/OSSTM.M.3.pdf>

Análisis del Marco Normativo Legal para el Ciclo de Vida de la Evidencia Digital

Mg. Aldo Santiago Igarza; Esp. Cintia Verónica Gioia; Mg. Jorge Eterovic;

Programa CyTMA2 / Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Florencio Varela 1903 (B1754JEC), San Justo, (5411) 4480-8900

asigarza@unlam.edu.ar; cgioia@unlam.edu.ar; eterovic@unlam.edu.ar

1. Resumen.

Con el crecimiento de las conductas delictivas que llegan a la justicia y que involucran dispositivos informáticos, surge la necesidad de acudir cada vez más a expertos en informática forense que actúen como peritos informáticos de oficio o de parte, siendo crucial su actuación en materia probatoria.

La exigente labor que hoy en día se requiere de especialistas en informática forense obliga a los mismos a mantener un conocimiento detallado y actualizado tanto a nivel de metodologías de prácticas forenses y procesos vinculados como en las normas y legislaciones asociadas con el tratamiento de la evidencia digital.

Disponer de un análisis comparativo exhaustivo de las diferentes normas y procesos del Ciclo de Vida de la evidencia digital dentro del marco normativo, jurídico y legal, son la base fundamental para el trabajo de un investigador forense informático o perito informático.

El proyecto de investigación se enfoca en el análisis de metodologías y procesos forenses informáticos y en el marco jurídico legal vigente para el aseguramiento del tratamiento de la evidencia digital en sus diferentes etapas del Ciclo de Vida, desde la identificación,

adquisición o recolección, preservación, análisis hasta la presentación de resultados técnicos a tribunales de la justicia.

Como resultado de esta investigación se espera poder hacer un aporte sobre cómo abordar un peritaje informático a partir de diferentes escenarios según el tipo de pericias a realizar, los tipos de delitos informáticos involucrados y la aplicación de la normativa legal vigente.

Palabras Clave:

Informática Forense, Evidencia Digital, Procesos Forense, Marco Normativo Legal.

2. Contexto.

Este proyecto de investigación está siendo presentado como un Programa de Investigación Científica, Desarrollo y Transferencia de Tecnologías e Innovaciones (CyTMA2) en el Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza.

El presente proyecto es del tipo investigación básica basado en el análisis del marco normativo y jurídico de la República Argentina, orientado a la comparación de los procesos del Ciclo de Vida de la Evidencia Digital.

3. Introducción.

La informática forense hace su aparición como una disciplina auxiliar de la justicia, para enfrentar los desafíos y técnicas de los intrusos informáticos, como también de garante de la verdad alrededor de la evidencia digital que se pudiera aportar en un proceso que involucra un delito informático [1].

De acuerdo con Dupuy y Kiefer [2], la evidencia digital es: “Cualquier información que, sujeta a una intervención humana u otra semejante, ha sido extraída de un medio informático”. La evidencia digital es un término utilizado para describir cualquier registro generado por o almacenado en un sistema computacional que puede ser utilizado como evidencia en un proceso legal [1].

La evidencia informática está comprendida por aquellos datos o información que se almacena, transmite o recibe en un dispositivo informático y que tiene valor probatorio en el marco de una causa judicial. La evidencia digital es intangible, inmaterial, volátil, frágil, anónima, duplicable, editable, ocultable y eliminable [3].

La informática forense es una disciplina científica técnica-legal que involucra el proceso de identificar, preservar, analizar y presentar evidencia digital, de manera que esta sea legalmente aceptable en una causa judicial [4].

La identificación de la evidencia digital implica identificar las fuentes potenciales de la evidencia digital, los elementos a secuestrar y documentar todo lo necesario para la identificación de esta.

La preservación de la evidencia digital consiste en extremar los recaudos a fin de evitar la contaminación de la prueba [5]. En esta etapa del proceso es muy importante mantener la cadena de

custodia registrando todas las operaciones realizadas sobre la evidencia digital y resguardando de forma segura los elementos secuestrados utilizando etiquetas de seguridad.

Los elementos de prueba originales deben ser conservados hasta la finalización del proceso judicial, preservándolos de las altas temperaturas, campos magnéticos y golpes. La creación de una imagen forense es indispensable y en el caso que no sea posible, el acceso a los dispositivos originales se realiza mediante mecanismos de protección contra escritura. Es importante implementar mecanismos de autenticación de la evidencia digital de manera de garantizar que la misma no fue alterada.

El análisis de la evidencia digital tiene como objetivo buscar y obtener evidencia digital relevante para la investigación (a partir de puntos de pericias solicitados), mediante la aplicación de diversas técnicas y herramientas forenses. Es importante registrar la evidencia digital relevante.

La presentación de la evidencia digital consiste en la elaboración del dictamen pericial con los resultados obtenidos en las etapas previas. El dictamen debe ser claro, objetivo y preciso, conteniendo la descripción de las tareas y elementos utilizados para repetir el proceso en caso de ser necesario [6].

Los resultados de un análisis científico de la evidencia digital deben poder ser repetibles, medibles e irrefutables.

La cadena de custodia involucra la custodia de todos los elementos del caso e incluye documentar cada uno de los eventos que se han realizado con la evidencia indicando por cada uno quién, cuándo, donde, en qué estado, quién tuvo acceso y adicionando toda información

que caracterice a como se llevó a cabo la custodia [7].

Los procesos relacionados con el manejo de la evidencia digital, asociados a la identificación, recolección, adquisición y preservación de datos, establecen principalmente cuáles son los requisitos para el manejo de la evidencia digital, los procedimientos a considerar para asegurar la cadena de custodia, los roles y responsabilidades del perito informático en cada etapa, la documentación a realizar y los componentes esenciales del Ciclo de Vida de la evidencia digital [8], [9].

4. Líneas de Investigación, Desarrollo e Innovación.

La informática forense es interdisciplinaria y requiere un estudio detallado de las normas, leyes, procesos, técnicas y tecnologías, además de los diferentes roles y responsabilidades de las personas involucradas, conformando un conjunto de conocimiento formal, científico y legal que apoya directamente a la administración de la justicia para al esclarecimiento de los hechos como así también en investigaciones internas en las organizaciones.

Si bien las herramientas forenses son la base esencial del análisis de la evidencia digital en medios informáticos, las mismas no hacen por sí solas a la tarea del perito informático. Por tal motivo el proyecto no se centra en la investigación de las herramientas forenses en sí, sino en la investigación de metodologías, técnicas, prácticas y procedimientos forenses y en el marco jurídico legal vigente para el aseguramiento del tratamiento válido de la evidencia digital en sus diferentes etapas del Ciclo de Vida.

También se estudiarán las diferentes regulaciones y lineamientos generales a

considerar para la implementación de un laboratorio de informática forense, de manera de basar la misma en un entorno regulado y basado en normativas de trabajo para la investigación forense.

5. Resultados y Objetivos.

El objetivo de este proyecto de investigación es abordar un análisis comparativo exhaustivo de las diferentes metodologías, procesos, procedimientos, normas, y prácticas relacionadas con la informática forense, dentro de un marco jurídico y legal como base para el trabajo del investigador forense informático o perito informático.

Se definirán los procesos relacionados con el manejo de la evidencia digital, asociados a la identificación, recolección, adquisición y preservación de datos, estableciendo principalmente cuáles son los requisitos para el manejo de la evidencia digital, los procedimientos a considerar para asegurar la cadena de custodia, los roles y responsabilidades del perito informático en cada etapa, la documentación a realizar y los componentes esenciales del Ciclo de Vida de la evidencia digital, dentro del marco normativo, jurídico y legal de nuestro país.

6. Formación de Recursos Humanos.

El equipo está integrado por docentes-investigadores que pertenecen distintas cátedras de la carrera de Ingeniería en Informática de la UNLaM, más otro docente-investigador abogado, especializado en temas jurídico-informáticos y un alumno de la carrera de Ingeniería en Informática que está haciendo sus primeras experiencias en investigación.

Dos de los miembros del equipo de investigación se encuentran desarrollando sus respectivos trabajos de tesis de posgrado de la Maestría en Informática de la UNLaM y en la Maestría en Teleinformática y Redes de Computadoras de la Universidad de Morón. Ambos están siendo tutorados por el Mag. Jorge Eterovic, integrante del proyecto de investigación.

7. Referencias

- [1] Cano, Jeimy J. *Computación Forense. Descubriendo los Rastros Informáticos*. 2da edición. Editorial Alfaomega. Bogotá, Buenos Aires, México. 2016. ISBN: 978-958-682-922-9.
- [2] Dupuy, D., Kiefer, M. “Ciberdelitos. Aspectos de Derecho penal y procesal penal. Cooperación Internacional. Recolección de evidencia digital. Responsabilidad de los proveedores de servicios de Internet”. Editorial B de F. Buenos Aires - Montevideo. 2017. ISBN: 978-9974-745-06-3.
- [3] Darahuge, M. E., Arellano González, L. “Manual de Informática Forense (Prueba Indiciaria Informático Forense)”. Editorial Errepar. Buenos Aires. 2011. ISBN: 978-987-01-1249-5.
- [4] Darahuge, M. E., Arellano González, L. “Manual de Informática Forense III (Prueba Indiciaria Informático Forense)”. Editorial Errepar. Buenos Aires. 2016. ISBN: 978-987-01-1953-1.
- [5] Sain, G., Azzolin, H. “Delitos Informáticos. Investigación criminal marco legal y peritaje”. Editorial B de F. Buenos Aires - Montevideo. 2017. ISBN: 978-9974-745-27-8.
- [6] Aboso, G. E., Zapata, M. F. “Ciberdelitos y Derecho Penal”. Editorial B de F. Buenos Aires - Montevideo. 2006. ISBN: 9974-578-74-4.
- [7] Marqués Arpa, T., Serra Ruiz, J. “Cadena de Custodia en el Análisis Forense. Implementación de un Marco de Gestión de la Evidencia Digital”. RECSI 2014, Alicante. España. Sitio web: <http://web.ua.es/en/recsi/2014/documentos/papers/cadena-de-custodia-en-el-analisis-forense-implementacion-de-un-marco-de-gestion-de-la-evidencia-digital.pdf> (visitado en marzo de 2018).
- [8] Domínguez, F.L. “Introducción a la Informática Forense”. Editorial RAMA. Buenos Aires. 2015. ISBN: 978-8499-642-09-3.
- [9] Di Iorio, A. “La necesidad de adopción de un Proceso Unificado de Recuperación de Información: “PURI – Una propuesta”. Congreso Argentino de Ingeniería Forense 2014. Copitec. Argentina. 2014. Sitio Web: <http://www.copitec.org.ar/comunicados/CAIF/2014/calidadserviciopericial.pdf> (visitado en marzo de 2018).

Software Abierto para la Evaluación de Sistemas Criptológicos Integrados

Cipriano, Marcelo¹; Malvacio, Eduardo; Estevez², Carlos; Fernández, Darío
García, Edith¹, López, Gabriel; Liporace, Julio¹; Maiorano, Ariel¹
Vera Batista, Fernando¹

¹Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática.
Escuela Superior Técnica, Facultad del Ejército. Universidad de la Defensa Nacional - UNDEF

²Instituto de Investigaciones Científicas y Técnicas para la Defensa – CITEDEF.-)

{marcelocipriano; dfernandez; verabatista}@est.iue.edu.ar
{edithgarcia; jcliporace; maiorano; edumalvacio}@gmail.com

RESUMEN

Es permanente la evaluación de la seguridad que debe llevarse a cabo en los sistemas de información, en la actualidad. Facilitar esta tarea tiene un alto impacto, así sea en la calidad de la seguridad del sistema, como en el aspecto económico de la misma al aumentar las herramientas de gestión de la seguridad, como así también un aporte a la reducción del tiempo de evaluación.

Este proyecto persigue el diseño y desarrollo de una herramienta que permita llevar adelante la automatización de tales análisis y la realización de pruebas que permitan realizar el estudio de manera veloz y eficiente.

Los usos de esta herramienta tiene aplicaciones DUAL, es decir puede destinarse para usos en el ámbito militar como también en el civil para ser aplicado en sectores gubernamentales, empresariales, educativos o privados.

La modalidad de Código Abierto o FOSS (por siglas en inglés de Free Open Source Software) permite y facilita la amplia difusión de los usos y aplicaciones de esta herramienta.

La evaluación de la seguridad implementada en un determinado sistema informático se ve ayudada por esta herramienta al permitir análisis estadístico de secuencias binarias para aplicar en algoritmos de Cifrado de Flujo (Stream Ciphers) y Generadores Pseudaleatorios

de Números (Pseudo Random Numbers Generators) y de la Complejidad Lineal. Tales secuencias pueden ser generadas por LFSRs (Linear Feedback Shift Registers), NLFSRs (Non-Linear Feedback Shift Registers), CCGs (Clock Controlled Generators), protocolos de seguridad de la información, programas o dispositivos para la Generación de Claves, Block Ciphers y demás algoritmos aplicados en entornos de Software como de Hardware.

Palabras Clave

Criptología, Criptoanálisis. Stream Ciphers.

CONTEXTO

En el marco de la carrera de grado de Ingeniería en Informática y el posgrado en Criptografía y Seguridad Teleinformática que se dictan en la *Escuela Superior Técnica "Gral. Div. Manuel N. Savio" (EST)*, dependiente de la *Facultad del Ejército, Universidad de la Defensa Nacional (UNDEF)* se llevan adelante tareas de I+D+i por parte del *Grupo de Investigación en Criptología y Seguridad Informática (GICSI)*.

GICSI depende del *Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática (Cripto-Lab)* perteneciente al *Laboratorio Informática (InforLab)*. Y está conformado por docentes investigadores, profesionales técnicos y alumnos de dicha área.

Esta línea de investigación aquí presentada corresponde a un proyecto que fue presentado, evaluado, aprobado y financiado en su integridad por la Universidad de la Defensa Nacional a través de su Programa UNDEFI, cuyos objetivos son:

- Fomentar estrategias colaborativas que incentiven la cooperación e integración transversal de las Facultades en investigación.
- Reconocer y potenciar capacidades de investigación existentes favoreciendo la interrelación entre grupos.
- Profundizar los esfuerzos institucionales dirigidos a la formación y perfeccionamiento de los recursos humanos para la investigación

1. INTRODUCCIÓN

Las comunicaciones del siglo XXI precisan de más y mejores algoritmos de cifrado, de autenticación y de integridad de la información que posibiliten la confidencialidad e integridad de las comunicaciones.

El aumento en los últimos tiempos en el empleo de Internet en general y el advenimiento en particular de las tecnologías de VoIP (voz sobre IP), Teleconferencias, VideoStreaming, VideoOnDemand y los Sistemas Móviles han incrementado la necesidad de asegurar estos servicios.

Los mecanismos de seguridad deben satisfacer las necesidades particulares de cada una de estas tecnologías, como por ejemplo las demandas de altas Tasas de Transmisión de Información, entornos de trabajo con recursos limitados de cálculo, memoria, energía y espacio, entre otros.

Las modernas técnicas criptológicas se han ido adaptando a estas necesidades que dieron nacimiento a llamada Criptografía Ligera o Liviana[1]. Es por ello que desde hace varios años muchos han sido los algoritmos que investigadores, universidades y empresas han dado

a conocer a la comunidad científica y al público en general.

El estudio y análisis de las propiedades matemáticas de los Generadores de Secuencias Pseudorandom generadas por:

- Substitution Boxes (S-BOXs) creadas para los modernos Block Ciphers[2].
- Stream Ciphers -en particular a aquellos algoritmos que involucran LFSRs (Linear Feedback Shift Registers).[3]
- NLFSRs (Non Linear Feedback Shift Registers)[4-6].
- CCG (Clock Controlled Generators)
- Autómatas Celulares
- Generadores de Números Pseudoaleatorios permite evaluar sus propiedades criptológicas a fin de poder identificar el nivel de robustez y seguridad de dichos algoritmos.

Muchas de las propiedades matemáticas y específicamente criptológicas se encuentran ocultas detrás del algoritmo. Esto quiere decir que al observar un algoritmo los autores describen con detalle el funcionamiento del mismo por medio de gráficos y demás especificaciones para “seguir la danza de los bits” y muchas veces las explicaciones técnicas más profundas no son abordadas tan en detalle y son parcialmente expuestas. Los investigadores deben profundizar en cada algoritmo y mediante su estudio, deducir sus propiedades.

A su vez en el ciclo de vida del algoritmo se encuentra una etapa que corresponde a los ataques a los que ha sido sometido. Muchas veces esos ataques son exitosos y van apareciendo vulnerabilidades. Otras veces es el estudio analítico y en detalle el que permite detectarlas. Muchas de ellas son resueltas con el advenimiento de versiones mejoradas de los mismos [7].

Es por ello que la permanente evolución de los algoritmos debe ir acompañada de la actualización de las pruebas y de-

más procesos que permitan evaluar la seguridad de los mismos.

Muchos algoritmos criptográficos fueron reconocidos por una norma internacional, como ser ISO (International Organization for Standardization) y IEC (International Electrotechnical Commission), conocidas como normas ISO/IEC:

- ISO/IEC18033-3:2005 llamada “Information Technology. Security Techniques. Encryption Algorithms. Part 3: Block Ciphers” Algoritmos de Cifrado en Bloques (Block Ciphers) de 64 bits: TDEA, MISTY1, CAST-128, HIGHT. De 128 bits: AES, Camellia, SEED. [8]
- ISO/IEC 29192-3:2012 “Information Technology. Security Techniques. Lightweight Cryptography. Part 3: Stream ciphers”. Algoritmos de Cifrado en Cadena (Stream Ciphers) Enocoro y Trivium.[9]
- ISO/IEC 18033-4:2011 “Information Technology. Security Techniques. Encryption Algorithms. Part 4: Stream Ciphers” presenta 5 algoritmos de Cifrado de Flujo: Decim-v2., KCipher-2 (K2), MUGI, Rabbit, SNOW 2.0. [10]

Aunque bienvenido, este renacimiento mundial por la búsqueda de nuevos algoritmos por sí sólo resulta insuficiente a la hora de establecer parámetros criptográficos seguros.

En la actualidad no existe una única modalidad general de criptoanálisis. Cada algoritmo, cada primitiva, cada protocolo debe ser atacado mediante una técnica adecuada a su estructura, si es que sus diseñadores atendieron a la necesidad de resistir los ataques conocidos.

El criptoanálisis tiene un impacto significativo en el mundo real, puesto que los algoritmos criptológicos, los protocolos y también los tamaños de las claves entre otros, son seleccionados ba-

sándose en el estado del arte del criptoanálisis.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Se ha dado en planificar este proyecto de investigación siguiendo 5 etapas:

- a) Investigación bibliográfica acerca de las técnicas de ataque criptoanalíticas conocidas y más difundidas.
- b) Estudio de factibilidad acerca del grado de automatización que requieren.
- c) Determinación de la metodología a automatizar.
- d) Elaboración de los módulos del programa y sus pruebas.
- e) Integración de los módulos en un marco general.

Los indicadores de avance previstos para el proyecto son:

- a) Identificación y selección de las herramientas matemáticas para la elaboración del software.
- b) Desarrollo de los módulos respectivos.
- c) Pruebas de los módulos.
- d) Integración en la aplicación final.

3. RESULTADOS OBTENIDOS / ESPERADOS

Se espera diseñar y desarrollar un software abierto que permita la evaluación de las propiedades criptográficas y de seguridad de secuencias pseudoaleatorias binarias procedentes de algoritmos criptográficos del tipo Stream Ciphers, Block Ciphers o Generadores de Números Seudoaleatorios.

El enfoque propuesto para el desarrollo se centra en el estudio e implementación de las diferentes técnicas criptoanalíticas aplicables de acuerdo a la bibliografía, su conjunción y el desarrollo de un conjunto de herramientas que permitan analizar las propiedades criptológicas.

Más allá del conocimiento que el equipo investigador obtenga de la realiza-

ción de este proyecto, se espera que el software que se desarrolle pueda ser empleado por:

- Administradores de Red que deseen probar la afectividad y seguridad de algoritmos criptográficos para su posterior uso.
- Investigadores que trabajen sobre vulnerabilidades en criptosistemas
- Organismos de Seguridad Informática.
- Docentes del área de Seguridad Informática y Criptografía para que puedan anexar esta herramienta al dictado de sus respectivas cátedras.

4. FORMACIÓN DE RECURSOS HUMANOS

Los docentes investigadores participantes del proyecto dictan las asignaturas *Criptografía y Seguridad Teleinformática, Matemática Discreta y Paradigmas de Programación I, II*. Desde esas cátedras se invita a los alumnos a participar en los proyectos de investigación. Es por ello que los alumnos *Cabrera Ezequiel* y *Dorado Mariano* forman parte del proyecto y próximamente se sumarán 2 alumnos, cuando UNDEF apruebe su incorporación. En particular los tres últimos serán postulantes para la beca “Estímulo a las Vocaciones Científicas” (EVC) otorgadas por el Consejo Interuniversitario Nacional (CIN) por encuadrarse en las condiciones pedidas. Asimismo este proyecto tiene una importante impronta para la formación de recursos humanos pues a los 6 investigadores con experiencia se les suman 2 investigadores que recién inician su experiencia.

Atendiendo a la responsabilidad ética y social que compete a la actividad científica y tecnológica, el Grupo Integrante de este Proyecto de Investigación, ya sea durante su ejecución o por la aplicación de los resultados obtenidos, desea expresar su compromiso a no realizar cualquier actividad personal o colectiva que pudiera afectar los derechos huma-

nos, o ser causa de un eventual daño al medio ambiente, a los animales y/o a las generaciones futuras.

5. BIBLIOGRAFÍA

- [1] <https://www.nist.gov/programs-projects/lightweight-cryptography> consultada el 10-3-18.
- [2] Daemen, J.; Rijmen, V.; *The Design of Rijndael: AES - The Advanced Encryption Standard*. Springer. New York. 2002.
- [3] Muller F., Peyrin T. *Linear Cryptanalysis of the TSC Family of Stream Ciphers*. Roy B. (eds.) *Advances in Cryptology - ASIACRYPT 2007*. Lecture Notes in Computer Science, vol. 3788. Springer, Berlin, Heidelberg. 2005.
- [4] Wu H., Preneel B. *Differential Cryptanalysis of the Stream Ciphers Py, Py6 and Pypy*. Naor M. (eds.) *Advances in Cryptology. EUROCRYPT 2007*. Lecture Notes in Computer Science, vol. 4515. Springer Berlin, Heidelberg. 2007.
- [5] Pasalic, E.; *On Guess and Determine Cryptanalysis of LFSR-Based Stream Ciphers*; IEEE Transactions on Information Theory. Vol.55 Ed.7º, 2009.
- [6] Ding C.; *The differential cryptanalysis and design of natural stream ciphers*. In: Anderson R. (eds.) *Fast Software Encryption. FSE 1993*. Lecture Notes in Computer Science, vol. 809. Springer Berlin, Heidelberg.
- [7] Dinur I., Shamir A. *Cube Attacks on Tweakable Black Box Polynomials*. *Advances in Cryptology - EUROCRYPT 2009*. Lecture Notes in Computer Science, vol 5479. Springer, Berlin, Heidelberg. 2009.
- [8] <https://www.iso.org/standard/37972.html> consultada el 12-3-18.
- [9] <https://www.iso.org/standard/59948.html> consultada el 12-3-18.
- [10] <https://www.iso.org/standard/54532.html> consultada el 12-3-18.

Técnicas y Herramientas para Regular la Seguridad en Web Services Basados en WSDL

Hernán Bernardis⁽¹⁾, Edgardo Bernardis⁽¹⁾, Mario M. Berón⁽¹⁾, Daniel E. Riesco⁽¹⁾, Maria Joao V. Pereira⁽²⁾

⁽¹⁾Departamento de Informática / Facultad Ciencias Físico Matemáticas y Naturales/ Universidad Nacional de San Luis
Ejército de los Andes 950 – San Luis – Argentina
{hbernardis, ebernardis, mberon, driesco}@unsl.edu.ar

⁽²⁾Departamento de Informática e Comunicações/ Instituto Politécnico de Bragança
Bragança - Portugal
mjoao@ipb.pt

Resumen

El desarrollo de sistemas en la actualidad ha mutado, siguiendo la tendencia mundial de migración hacia la nube. Se logra mayor escalabilidad al construir sistemas con módulos distribuidos en la red, en donde el sistema final es la combinación de un conjunto de módulos distribuidos en la nube. Este concepto ha funcionado a lo largo de la historia por medio de diferentes arquitecturas tecnológicas - RPC, RMI, etc. - pero, desde hace algunos años, los servicios web han sido la arquitectura de mayor popularidad. Su independencia de la arquitectura tecnológica subyacente junto con su aprovechamiento de los protocolos ya definidos de internet son algunos de los motivos impulsores de su popularidad.

En el caso particular de esta investigación, el objeto de estudio son los servicios web con descripciones en WSDL. La idea de este trabajo consiste en detectar y - de ser necesario - modificar el nivel de entendimiento que posee la descripción de un servicio web para minimizar las vulnerabilidades que este pueda tener al estar publicado en la nube. Para detectar estas vulnerabilidades, es primordial primero comprender el servicio web y, por ende, su descripción.

En este artículo se describe una línea de investigación centrada en facilitar la comprensión de Servicios Web mediante el análisis de sus especificaciones WSDL. Este análisis sirve como base para el cálculo del grado de entendimiento del mismo y, en base a esta medida, determinar qué acciones se deben tomar, en caso de que corresponda, para cambiar su grado de entendimiento ocultando vulnerabilidades.

Palabras clave: Web Services, Métricas, WSDL, Comprensión, Seguridad.

Contexto

La línea de investigación descrita en este artículo se desarrolla en el Laboratorio de Calidad e Ingeniería de Software (**LaCIS**) de la Universidad Nacional de San Luis (UNSL); y se encuentra enmarcada dentro de dos proyectos. “*Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de software de Calidad*”, perteneciente a la UNSL. Proyecto reconocido por el programa de incentivos y continuación de diferentes proyectos de investigación de gran éxito a nivel nacional e internacional. También dentro del proyecto bilateral con la Universidade do Minho (Portugal): “*Fortalecimiento de la Seguridad de los Sistemas de Software*”

mediante el uso de *Métodos, Técnicas y Herramientas de Ingeniería Reversa*”, recientemente aprobado por el Ministerio de Ciencia Tecnología e Innovación Productiva (Mincyt) y su código es PO/16/93.

Introducción

Actualmente, con el auge de internet se están popularizando los Web Services como artefactos de software a partir de los cuales se pueden construir sistemas más complejos. Haciendo una burda comparación, los web services son los ladrillitos Lego[®] de la programación, que al unirse permiten crear sistemas mayores y muy potentes.

Según la W3C, un Web Service es: *“Una aplicación de software identificada por una URI, cuya interface y enlaces son capaces de ser definidos, descriptos y descubiertos como artefactos XML. Un web service soporta interacción directa con otros agentes de software usando mensajes basados en XML intercambiados a través de protocolos basados en internet”*. Muchas organizaciones construyen sus sistemas basándose en una arquitectura orientada a servicios web, en donde algunos de ellos se publican al resto del mundo y otros simplemente son utilizados de manera interna por sus equipos de desarrollo. Esta descentralización permite que cada equipo de desarrollo elija la arquitectura que desee para construir sus proyectos sin afectar la vinculación con el resto del sistema. La interacción entre proyectos se convierte en un intercambio de mensajes con la información necesaria dentro, sin necesidad de vinculación a nivel de arquitectura subyacente más que la necesaria a la invocación de los servicios web. Es justamente esta vinculación la que planteó el desafío más complejo en el mundo de los servicios web: de qué manera lograr quitar la arquitectura tecnológica subyacente de los sistemas para que la interacción no se viera

afectada?. La creación de protocolos específicos de descripción junto al uso de protocolos de internet fueron la solución.

La idea de construir un Web Service y que pueda ser utilizado por cualquier otra persona u organización en el mundo ha sido posible debido a la creación de estándares y lenguajes formales de definición de los mismos. Sin embargo, esta alta abstracción en la construcción y especificación de los mismos dificulta en gran medida su comprensión. Comprensión necesaria tanto a la hora de realizar tareas de mantenimiento (adaptación, arreglo de errores, migración, etc.) como a la hora de analizar las vulnerabilidades que estos servicios pueden dejar al descubierto hacia el mundo. Es importante recordar que un servicio web es una “interfaz” al mundo de un sistema de software que puede ser atacado y vulnerado.

Todo Web Service posee una especificación que provee la información necesaria para invocarlo. Uno de los estándares de descripción más conocido es WSDL (Web Service Definition Language) [1]. Las especificaciones WSDL son un dialecto XML, con reglas bien definidas para especificar cada componente del WS. Cuántos parámetros recibe y de qué tipo son, qué datos retorna y de qué tipo, qué protocolo de internet usa para su comunicación, qué operaciones posee, son entre otras tantas, características del WS que se encuentran especificadas en su WSDL asociado.

Así como el archivo WSDL sirve para que un agente de software o persona pueda interpretarlo para usar el servicio web que describe, también puede dar información a personas no deseadas o incluso exponer vulnerabilidades. Más aún si se considera que existen herramientas que generan los WSDLs de manera automática para un servicio web, con lo cual el nivel de atención a la información que se publica no siempre se encuentra bajo un estricto control. Esto se vuelve más importante

para aquellos casos en donde los servicios web pertenecen a bancos, tarjetas de créditos, servicios de compra/venta online, entre otros. Incluso también para los servicios web que no se publican, son privados y necesitan mayor control y seguridad como los que pertenecen a empresas privadas y redes militares.

Empresas competidoras pueden aprender el know-how y conseguir copiar el diseño para ofrecer servicios similares y competitivos. Pero no solo se trata de competencia, los ataques de seguridad como espionaje de información, suplantación de clientes, inyección de comandos y denegación de servicio también son posibles ya que los atacantes pueden aprender sobre los datos intercambiados y los patrones de invocación de los documentos WSDL. Si bien la legibilidad de las descripciones de los servicios hace que los servicios web sean reconocibles, también contribuye a la vulnerabilidad del servicio [2].

Suena lógico entonces analizar la seguridad que posee un determinado WSDL para controlar la seguridad del mismo. Para determinar esto, es necesario primero comprender el SW mediante el análisis de su correspondiente especificación WSDL.

Aprovechando la alta estandarización presente en el lenguaje WSDL, se pueden definir metodologías de comprensión de los mismos mediante la extracción y análisis de la información presente en dichas especificaciones [3, 4, 5]. A partir de esta información, determinar qué tan entendible es un WSDL y qué vulnerabilidades puede reflejar. Luego, usando diferentes técnicas, se pueden disminuir estas vulnerabilidades, ocultándolas o, en casos extremos, forzando la reconstrucción del WSDL o su SW asociado [6].

En este trabajo se extrae información aplicando técnicas de compilación, algoritmos

de análisis de lenguaje natural y técnicas de cálculo de indicadores sobre su especificación WSDL. Toda esta información se utiliza para, a partir del cálculo de métricas propias, determinar la dificultad de comprensión que poseen [7, 8]. Además, también se utiliza LSP (Logic Scoring of Preference) para definir estructuras de agregación que le asignen pesos a los valores de cada métrica según sea la necesidad del ingeniero de software y, en base a estos pesos, se obtiene un grado de entendimiento global de la especificación WSDL [9,10]. Toda esta información se usa de base para definir qué partes del WSDL muestran vulnerabilidades y, qué modificaciones se pueden realizar sobre el mismo para mitigar esto.

La organización de este artículo se expone a continuación. La sección 2 describe la línea de investigación abordada. La sección 3 presenta los resultados obtenidos hasta el momento, junto con aquellos esperados a corto plazo. Finalmente, la sección 4 describe las tareas realizadas por los recursos humanos en formación.

Líneas de Investigación y Desarrollo

El análisis y reducción de vulnerabilidades de las especificaciones WSDL posee múltiples etapas con su función particular dentro del proceso global. En las subsecciones siguientes se describen brevemente dichas etapas.

Extracción de Información

Debido a que las especificaciones WSDL son un dialecto XML, se pueden usar técnicas de compilación sobre las mismas basadas en los parser DOM (Domain Object Model) [11]. Un parser DOM construye una representación interna del WSDL basada en árboles. A partir de funciones específicamente diseñadas para recorrer la representación construida (funciones

transversales) se extrae la información deseada. Estas funciones transversales logran extraer los identificadores de cada componente (*name, type element, etc.*), la documentación y los comentarios presentes en el WSDL.

Cálculo de Métricas

De la información extraída del WSDL, se calculan múltiples métricas, que pueden ser lógicamente diferenciadas en los siguientes grupos:

- **Métricas de tamaño:** miden la complejidad del WSDL en base a las cantidades de componentes de cada etiqueta dentro del WSDL, como por ejemplo cantidad de tipos complejos, de parámetros, de operaciones, de mensajes, entre otras. Esto permite tener una idea del tamaño de cada sección particular del WSDL y determinar qué tan complejo es, a primera vista, su comprensión.
- **Métricas de calidad:** permiten medir la calidad semántica de la especificación WSDL. Esto es, que tanta información semántica brinda la especificación WSDL respecto del WS que representa y que tan entendible y comprensible en sí es dicha especificación.
- **Métrica de entendimiento global:** usando LSP se calcula el grado de entendimiento que posee la especificación WSDL de un WS [9, 10].

Incrementar la Seguridad

Toda aplicación web está conformada por distintos tipos de información, tanto formal como informal. El análisis detallado de la misma y el cálculo de métricas permite detectar, dado su grado de entendimiento, que partes son más susceptibles a los ataques. En este punto, es posible definir estrategias que permitan subsanar las vulnerabilidades y proteger las partes que sean susceptibles de ataques [6].

Utilizando la información extraída del WSDL se pueden manipular diferentes partes del mismo para mejorar su seguridad disminuyendo su nivel de entendimiento. Esto se puede lograr mediante la utilización de funciones de ofuscación y/o encriptación al realizar las modificaciones y/o transformaciones necesarias que aumentaran el nivel de seguridad. Estas transformaciones pueden ser sobre partes específicas del WSDL (identificadores, operaciones, etc.) o en la totalidad del mismo. Dichas modificaciones dependen del nivel de seguridad deseado, partiendo de un nivel básico en donde se ofuscan y/o encriptan partes específicas del WSDL, como por ejemplo el nombre de los identificadores, hasta llegar a un nivel máximo en donde se realiza una transformación completa del WSDL.

Resultados Obtenidos/Esperados

Algunos de los resultados más destacados obtenidos por esta investigación son:

- Se definieron y calcularon diferentes métricas (cuantitativas y cualitativas) que permiten medir la complejidad de los WS.
- Se utilizó LSP para calcular el grado de entendimiento global del WS.
- Se construyó WSDLUD, una herramienta que automatiza el proceso de cálculo de métricas, del grado de entendimiento del WS usando LSP y la visualización de la información.
- Se definieron métodos y herramientas para la reducción de vulnerabilidades del WSDL mediante su ofuscación y/o encriptación.

Entre los objetivos planteados a corto y largo plazo se pueden mencionar:

- Mejorar las técnicas de mitigación de vulnerabilidades.
- Construir una herramienta que, vinculada con WSDLUD, permite realizar la transformaciones de ofuscación/criptación de manera automática.
- Ampliar y aplicar el prototipo a especificaciones escritas en BPEL debido a que este lenguaje es utilizado para la ejecución de procesos de negocios.
- Estudiar, comprender y ampliar el número de métodos de encriptación y ofuscación de código utilizados.

Formación de Recursos Humanos

Las tareas realizadas en el contexto de la presente línea de investigación están siendo desarrolladas como parte de trabajos para optar al grado de Magister en Ingeniería de Software. En el futuro se piensa generar diferentes tesis de maestría y doctorado a partir de los resultados obtenidos de este trabajo.

Bibliografía

- [1] WSDL Specification for W3C <https://www.w3.org/TR/wsdl>.
- [2] Pananya Sripairojthikoon, Twittie Senivongse. "Concept-Based Readability Measurement and Adjustment for Web Services Descriptions". ICACT Transactions on Advanced Communications Technology (TACT) Vol. 3, Issue 1, January 2014.
- [3] N. Gold and K. Bennett. "Program comprehension for web services". In Program Comprehension, 2004. Proc. 12th IEEE International Workshop on. June 2004.
- [4] L. O'Brien Lero and D. Smith. "Working session: program comprehension strategies for web service and service oriented architectures". Proc. of 12th IEEE International Workshop on Program Comprehension. 2004.
- [5] H. El Bouhissi, M. Malki, and D. Bouchiha. "A reverse engineering approach for the web service modeling ontology specifications". In Sensor Technologies and Applications 2008. SENSORCOMM '08. Second International Conference on, pages 819–823, Aug 2008.
- [6] Edgardo Bernardis, Hernán Bernardis, Mario Berón, Germán Montejano. "Seguridad en Servicios Web". XIX Workshop de Informática y Ciencias de la Computación (WICC). Buenos Aires, Argentina. Abril de 2017.
- [7] C. Mao. "Towards a data complexity metric set for web service composition". In Computer and Information Technology (CIT), 2011 IEEE 11th International Conference on, pages 127–131, Aug 2011.
- [8] Fangfang Liu, Yuliang Shi, Jie Yu, Tianhong Wang, Jingzhe Wu. "Measuring Similarity of Web Services Based on WSDL". IEEE International Conference on Web Services ICWS. 2010.
- [9] Bernardis, Hernán; Berón Mario; Bernardis, Edgardo; Riesco, Daniel; Henriques, Pedro. "Extracción de información y cálculo de métricas en WSDL 1.1 y 2.0". II Congreso Nacional de Ingeniería Informática / Sistemas de información (CoNaIISI). Argentina. 2014.
- [10] Mario M. Berón, Hernán Bernardis, Enrique A. Miranda, Daniel E. Riesco, Maria João Pereira, Pedro Rangel Henriques. "WSDLUD: a Metric to Measure the Understanding Degree of WSDL Descriptions". Proceedings of the 2015 Symposium on Languages, Applications and Technologies, SLATE'15. Madrid, España 2015.
- [11] Parser DOM specification for W3C. <https://www.w3.org/DOM>.

Investigación en ciberseguridad: un enfoque integrado para la formación de recursos de alto grado de especialización

Javier Díaz, Lía Molinari, Paula Venosa, Nicolás Macia, Einar Lanfranco, Alejandro Sabolansky

Laboratorio de Investigación de Nuevas Tecnologías Informáticas (LINTI). Facultad de Informática. Universidad Nacional de La Plata. 50 y 120. La Plata

{javierd,lmolinari,pvenosa,nmacia,einar,asabolansky}@info.unlp.edu.ar

Resumen

Desde los inicios de la década del 2000, el LINTI desarrolla una línea de investigación en seguridad, inicialmente, y ciberseguridad en la actualidad, que no ha tenido interrupciones. Importantes experiencias como CERTUNLP, PKIGRID UNLP, entre otras, fueron la implementación que ejercitan las actividades de investigación que se realizan. La evolución de las TICs conlleva situaciones de riesgo que se van revelando día a día. Tecnologías emergentes habilitan el tratamiento de gran cantidad de datos, pero también habilitan su exposición. La investigación sobre tendencias y comportamientos es una tarea de investigación que muchas veces trasciende lo informático. En este artículo se afianza la línea de investigación existente con la inclusión de nuevos desafíos como IoT, forensia, desarrollo seguro de aplicaciones, normativa, estándares y buenas prácticas, y el habitual compromiso de trasladar los resultados de la investigación a la docencia y a la extensión, enfocados a formar profesionales con habilidades, capacidades y conocimientos para entender y administrar adecuadamente los sistemas de gestión de la ciberseguridad.

Palabras clave: seguridad de la información, CSIRTs, Internet de las cosas, forensia, PKI

Contexto

La línea de investigación “Ciberseguridad” presentada en este trabajo, se inserta en el proyecto de investigación "Internet del Futuro: Ciudades Digitales Inclusivas, Innovadoras y Sustentables, IoT, Ciberseguridad, Espacios de Aprendizaje del Futuro" del Programa Nacional de Incentivos a docentes-investigadores, que se desarrolla en el LINTI de la Facultad de Informática de la Universidad Nacional de La Plata (UNLP). Este proyecto está acreditado por la UNLP y financiado por partidas del presupuesto nacional.

Introducción

Las Tecnologías de la Información y las Comunicaciones (TIC) están incorporadas en todos los aspectos de la vida cotidiana. Entre otras situaciones, los precios cada vez más accesibles de los dispositivos de conexión a Internet han permitido su uso masivo, y han pasado a ser un accesorio imprescindible en la vida de las personas y las organizaciones. Ubicuo, instantáneo, actualizado, son adjetivos que acompañan el anuncio de los servicios tecnológicos.

No sólo el abaratamiento de los dispositivos ha contribuido a esa apropiación: la facilidad de uso, que no requiere un alto grado de especialización (el período entre la adquisición y la puesta en funcionamiento es cada vez más corto), la posibilidad de almacenamiento en la

nube terminó con las restricciones de espacio y la oferta de aplicaciones en diferentes “stores” invitan a la instalación entusiasta.

Ya no sólo se conectan personas y máquinas: Internet de las cosas (Internet of Things, IoT) habilita la conexión de las máquinas entre sí, sin la intervención de humanos.

No obstante, todas estas ventajas y beneficios pueden transformarse en una verdadera amenaza para quienes no tienen recaudos. Robo de identidad, grooming, carding, entre otras, son situaciones no deseadas que llegaron junto con las TICs.

La promulgación de leyes que contemplan la figura del delito informático, ha alimentado la esperanza que se condene a quienes lo cometen. Pero la determinación de responsabilidades en el mundo digital exige un conocimiento especializado y experiencia en una disciplina en constante evolución.

Ante una denuncia de un potencial delito, los diferentes actores que intervienen necesitan recurrir a profesionales con un alto grado de especialización no sólo desde el punto de vista técnico, sino legal y ético. Esto ha generado una demanda de capacidades y competencias que habitualmente no están consideradas en los contenidos de las currículas.

Si bien siempre la normativa surge luego del uso, un alto grado de sensibilización sobre las vulnerabilidades y amenazas en el mundo digital ha generado una importante cantidad de estándares, buenas prácticas y guías para la gestión de la ciberseguridad. Regulaciones y directivas sobre los datos, intentan establecer pautas claras acerca de la garantía de privacidad.

Al mismo tiempo, la enseñanza basada en competencias, tendencia que está presente en las Reformas educativas de fines de los 90 en algunos países como Francia, Bélgica, hoy ya ha alcanzado también a los países latinoamericanos (Spiegel, 2008) (IEEE, 2013). Estas formas de aprender se ponen en práctica en todos los niveles educativos y se aplican fácilmente en la educación superior.

El Laboratorio de Investigación en Nuevas

Tecnologías Informáticas (LINTI) de la Facultad de Informática de la Universidad Nacional de La Plata sostiene desde el año 2001 una línea de investigación, docencia, extensión y transferencia en ciberseguridad. La publicación de artículos en congresos nacionales e internacionales, la participación en competencias CTF (Capture the flag) con excelentes resultados, la creación de una Infraestructura de Clave Pública (Public Infrastructure Key, PKI) en el 2006¹ y su continuidad a la fecha, la creación y administración de CERT UNLP², la participación en proyectos de Extensión y Transferencia, y la actualización constante de los contenidos de las asignaturas a su cargo, entre otros logros, han permitido contar con un equipo de profesionales con alto grado de especialización tanto en conocimientos como en experiencia, sumado al perfil ético que garantizan una formación integral.

En este contexto de constante desafío se hace necesario incorporar contenidos relacionados con estas tecnologías emergentes y las situaciones de riesgo que conllevan, en docencia, investigación, extensión y transferencia.

Líneas de Investigación y Desarrollo

La ciberseguridad es una temática que es transversal a las diferentes ramas de las TICs. Redes, bases de datos, autenticación, desarrollo de aplicaciones, entre otros, deben dedicar un espacio a la ciberseguridad para poder hablar de “contextos seguros”.

La línea de investigación que se describe aborda, entre otros:

1. Tendencias en incidentes de seguridad y su gestión. Como integrantes del Centro de Respuesta a Incidentes de Seguridad de la Información de la Universidad Nacional de La Plata (CERTUNLP), a

¹ www.pkigrd.unlp.edu.ar

² www.cert.unlp.edu.ar

la hora de gestionar los incidentes resulta fundamental conocer cuáles son los problemas más críticos a los que las organizaciones se enfrentan hoy en día (Venosa P.,2014), las técnicas que se utilizan para llevar a cabo los ataques y cómo mitigar los mismos (Francisco Javier Díaz,2017).

2. Forensia digital. La forensia digital constituye una etapa fundamental en el proceso de gestión de incidentes a la hora de investigar las características de un incidente, permitiendo obtener los detalles de lo que ocurrió y aprender de ello. Resulta de gran importancia investigar las técnicas utilizadas en el análisis forense y seleccionar herramientas que den soporte para automatizar las tareas asociadas. (Diaz J., 2016)
3. Infraestructura de clave pública PKI. En el marco de Americas Grid Policy Management Authority³ (TAGPMA), la prestación del servicio de emisión de certificados digitales a través de PKI Grid UNLP, incluye el estudio de vulnerabilidades que surgen en relación al manejo de claves, algoritmos, etc, así como el análisis de los protocolos, procedimientos y herramientas que se utilizan a fin de garantizar que la gestión de certificados cumple con los requisitos de seguridad exigidos.
4. Seguridad en IoT. La aparición y evolución de IoT que trae consigo la posibilidad de conectar todo tipo de objeto doméstico a las redes, no ha requerido grandes cambios en la estructura de Internet, ya que varios protocolos del stack TCP/IP se reutilizan en IoT (IPv6, UDP, TCP, HTTP, entre otros). En este marco, es necesario que las aplicaciones aseguren la confidencialidad, integridad y autenticidad de los datos que almacenan

y se envían entre las diferentes componentes. En consecuencia, debe proveerse seguridad tanto en dichas componentes como en el proceso de comunicación entre las mismas, teniendo en cuenta las diferentes capas de red y la sensibilidad de los datos que las mismas manejan. (Diaz J., 2017)

5. Desarrollo seguro de software. Para entender cómo desarrollar de forma menos insegura es necesario conocer cuáles son los problemas de seguridad que afectan al software, cuáles son las consecuencias de que alguien explote esos problemas y cuáles son las mejores técnicas y herramientas disponibles para evitar desarrollar explotable.
6. Entrenamientos de seguridad. Los CTF (Capture the flag)(Vigna, 2014) son competencias de seguridad informática donde un equipo puede poner en práctica sus habilidades para descubrir vulnerabilidades, explotarlas así como también resolver problemas. Hay dos tipos de competencias:
 - CTF – Capture The Flag: Competencias de tipo ataque / defensa
 - Jeopardy: Competencias tipo pregunta / respuesta, donde se presentan distintas categorías de desafíos, como ser: forensia, esteganografía, criptografía, redes, explotación web, etc.
7. Normativa, marcos referenciales y buenas prácticas. Es sabido que la normativa se define luego del uso. Los estándares genéricos son la base para la elaboración de normativa en general, políticas y estrategias. Enunciadas tendencias acerca de regular la privacidad se hacen realidad como es el caso de European General Data Privacy Regulation (EU - GDPR)⁴, con fecha de

³ www.tagpma.org

⁴ <https://www.eugdpr.org/>

aplicación en Mayo del corriente año. Conceptos como anonimización (anonymization) y pseudo-anonimización (pseudo-anonymization) empiezan a preocupar a los administradores de big data. Evidencia del interés es esta línea de trabajo son varios artículos publicados (Díaz, Molinari, 2017).

Resultados obtenidos/esperados

Como principales objetivos se plantean:

- Consolidar la línea de investigación en ciberseguridad y su aplicación en la docencia y la extensión, trabajando sobre los temas emergentes asociados a las metodologías y paradigmas que surgen día a día.
- Promover las prácticas en lo que hace a tener en cuenta la seguridad en todas las etapas del ciclo de vida del desarrollo, de los servicios y de la gestión de la organización.
- Transmitir la experiencia adquirida en los distintos proyectos y actividades a los alumnos de las cátedras de grado y postgrado con contenidos afines de nuestra Facultad.
- Realizar actividades de difusión, capacitación y sensibilización en el uso del espacio digital con la comunidad, relacionados con las problemáticas actuales que los involucran.
- Conformar un equipo interdisciplinario para el análisis de situaciones que trascienden lo del mundo de las TICs, para evaluar tendencias y propuestas de sensibilización en el uso del espacio digital.

Partiendo de esos objetivos y dentro de las temáticas y proyectos descritos en secciones anteriores se ha arribado a los siguientes resultados:

- Como parte del proyecto vinculado al Centro de Excelencia en Ciberseguridad de la ITU5, del que formamos parte, se han diseñado

los cursos “Ciberseguridad: primeros pasos de un gran desafío” y “CSIRT: coordinando prevención, detección, manejo de incidentes, respuesta y mitigación de ciberataques”, que son parte de la oferta de formación para el año 2018. En mayo de 2018 se realizará el Cyberdrill6 de la ITU y la UNLP será anfitrión. El equipo de Ciberseguridad de la UNLP será el organizador.

- También a partir de la investigación en esta línea se han armado y dictado varias iniciativas en 2017: un taller de seguridad en IoT en el marco del curso de Postgrado “Internet de las cosas”; un taller de seguridad Informática en la UNNOBA, dirigido a docentes, graduados y alumnos avanzados; y se ha organizado una jornada de seguridad en la Facultad de Informática de la UNLP en el marco de la semana de la seguridad informática en noviembre de 2017.

- Desde el año 2017 se planificaron y realizaron encuentros semanales en los que un grupo de docentes y alumnos trabajan en temáticas relacionadas a las competencias de seguridad, intercambiando experiencias adquiridas en los últimos concursos en los que hubo participación y estudiando nuevas metodologías. Esta actividad se enmarca en los grupos de interés definidos por la Secretaría de Innovación de la Facultad.

Formación de Recursos Humanos

El equipo de trabajo que lleva adelante estas líneas forma parte del LINTI. El mismo está formado por docentes/investigadores de la Facultad de Informática de la UNLP. En el marco de sus actividades, tiene entre sus funciones la dirección de tesis relacionadas con las temáticas en las que trabaja.

Ese equipo ha dirigido varias tesinas de grado

https://academy.itu.int/index.php?option=com_content&view=article&id=154&Itemid=588&lang=en

⁶ <https://www.itu.int/en/ITU-D/Regional-Presence/Americas/Pages/EVENTS/2018/20577.aspx>

en la presente línea de investigación. Aquéllas que se han finalizado en el último año han sido: “Análisis y mejoras de seguridad a una aplicación prototipo en IoT” del alumno Bruno Pertini, “Análisis e implementación de RPZ” del alumno Matías Ferrigno, “Uso de smartphones para auditar la seguridad de redes inalámbricas” de los alumnos Juan Ignacio Bernal y Alejandro Zurita (Venosa P.,2017) y “Ampliación y mejora de servicios en la infraestructura de clave pública para e-ciencia de la UNLP (PKIGrid UNLP)”. de los alumnos Guido Celada y Juan Manuel Filandini.

Además como trabajo final de la cátedra Seguridad y Privacidad en Redes de la cual dos docentes de este grupo forman parte, dos alumnos han investigado sobre el protocolo LORA y el análisis de su seguridad, continuando con temáticas de seguridad en IoT a fin de fortalecer este eje de investigación.

En el marco del proceso de formación continua del equipo, el profesor Nicolás Macia, ha presentado su tesis para obtener el título de Magister en Redes de Datos de la UNLP, realizando el trabajo titulado “Diseño y desarrollo de un mecanismo más seguro de manejo de sesiones web”. Finalmente, la profesora Lía Molinari ha expuesto su tesis doctoral titulada "Modelo de Gestión para la Prevención de Lavado de Activos (PLA) en el sector de juegos de azar". Este trabajo plantea un enfoque orientado al riesgo basado en diferentes marcos referenciales y estándares en la gestión de la seguridad de los sistemas de información.

Bibliografía

IEEE (2013).
<http://ieeexplore.ieee.org/abstract/document/6521304?reload=true>

Venosa P., Díaz J. (2014). Detección de botnets utilizando herramientas Opensource. WICC 2014. ISBN 978-950-34-1084-4

Diaz J., Molinari L., Belalcazar A. Ron M (2017). Towards a Strategic Resilience of Applications through the NIST Cybersecurity Framework and the Strategic Alignment Model (SAM). INCISCOS 2017. Quito, Ecuador.

Javier Diaz, Paula Venosa, Nicolas Macia, Einar Lanfranco,Alejandro Sabolansky, Damian Rubio, (2016). Análisis digital forense utilizando herramientas de software libre. WICC 2016. ISBN: 978-950-698-377-2

Díaz J., Venosa P., Fava L., Castro N., Vilches D., López F. (2017). Estrategias de IoT para Lograr Ciudades Digitales Seguras, más Inclusivas y Sustentables. WICC 2017. ISBN 978-987-42-5143-5.

Venosa P., Diaz J. (2014). Detección de botnets utilizando herramientas Opensource. WICC 2014. ISBN 978-950-34-1084-4

Venosa P., Macia M., Lanfranco E., Sabolansky A. (2017). Análisis de Seguridad en Redes Wireless Utilizando Dispositivos Móviles. WICC 2017. ISBN 978-987-42-5143-5.

Stiefel, Berta Marco (2008). Competencias básicas: Hacia un nuevo paradigma educativo. Narcea Ediciones.

Francisco Javier Díaz, Alejandro Sabolansky, Nicolás Macia, Paula Venosa, Einar Lanfranco, Mitigación de DDoS en Redes Académicas e IXPs . CIBSI 2017.

Seguridad en la Nube: Almacenamiento de Imágenes Médicas Y Watermarking

Silvia Edith Arias¹, Laura Mónica Vargas^{2,3}, Alejandra Di Gionantonio¹, Daniel Arch¹, Diego Serrano¹, Martín Navarro Mugas¹, Nicolás Hernandez¹, Paula Sosa¹, Ezequiel Ambrogio¹

s_autn@hotmail.com, laura.monica.vargas, ing.alejandradg, diegojserrano, mnavarromugas, damiannicolas05, sosa.pau, ezequielambrogio@gmail.com, daniel.arch@pjn.gov.ar

¹Laboratorio de Investigación de Software, Departamento de Ingeniería en Sistemas de Información, Facultad Regional Córdoba, Universidad Tecnológica Nacional

²Laboratorio de Redes y Comunicaciones de Datos, Departamento de Computación, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba

³ Laboratorio de Procesamiento de Señales, Departamento de Matemática, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba

I. RESUMEN

En los últimos años, se desarrollaron muchas aplicaciones acompañando el rápido avance de las telecomunicaciones. Una de ellas es la telemedicina por medio de la cual los médicos pueden transferir y compartir los datos digitales de los pacientes en forma remota para determinar un diagnóstico definitivo. Actualmente, la información médica que se almacenaba en el centro de salud se lleva a la nube. Por lo tanto, es esencial proteger los datos médicos intercambiados, especialmente cuando se utiliza una plataforma de Cloud Computing donde la seguridad es un problema importante. Hay que garantizar que las imágenes médicas se puedan compartir en forma segura preservándolas de cualquier intento de distorsión. Como así también proporcionar privacidad en las cadenas de datos de los Registros Electrónicos de los Pacientes o Electronic Health Records (EHR).

En este trabajo, nos dedicamos a explorar y obtener conocimientos teóricos sobre marcos de trabajo de plataformas de Health Cloud Computing que permitan alojar imágenes médicas con inserción de Marcas de Agua en los EHRs. La finalidad de este trabajo de campo es seleccionar y probar plataformas de Health Cloud Computing. Posteriormente elaborar una comparación cuantitativa y cualitativa de sus características

principales, y recomendar cuáles son las plataformas más seguras y adecuadas.

Palabras clave: seguridad informática, watermarking, imágenes digitales, cloud computing, telemedicina.

II. CONTEXTO

El presente trabajo se realiza en el Laboratorio de Investigación de Software, Departamento de Ingeniería en Sistemas de Información de la Facultad Regional Córdoba de la Universidad Tecnológica Nacional. En el marco del Proyecto “Análisis comparativo entre Plataformas de Cloud Computing, para el caso de almacenamiento de imágenes médicas con marcas de agua” acreditado y financiado por la Secretaría de Ciencia y Técnica de Código: CCUTNCO0004961. El cual se lleva a cabo en el Laboratorio de Investigación de Software de la Facultad Regional Córdoba de la Universidad Tecnológica Nacional (Argentina). La temática de watermarking ha sido presentada por los cuatro primeros autores en los Proyectos homologados por SeCyT-UTN PID, “Marcas de Agua múltiples en imágenes digitales fijas para autenticación y detección de adulteraciones”. Código SCyT – UTN1166, 2010-2011, Resolución 26/10, 2010 SCyT del Rectorado de UTN y “Marcas de Agua Seguras en Imágenes para identificación del propietario”. Proyecto ID

promocional. Código SCyT- UTN EIPRCO753, 2008-2009, Resolución 75/08 SCyT del Rectorado UTN.

La segunda y la tercera autora han publicado en la Revista de la FCEFYN de la UNC un artículo de difusión de marcas de agua. "Marcas de Agua: una Contribución a la Seguridad de Archivos Digitales". Revista de la Facultad de Ciencias Exactas, Físicas y Naturales de la UNC. ISSN 2362-2539 (Versión electrónica). Año 3 – N° 1 (2016).

III. INTRODUCCIÓN

En la etapa actual, la investigación se centra en el análisis del problema de seguridad de la información contenida en las imágenes médicas cuando estas son alojadas en un framework de Cloud Computing.

Cloud Computing es un mecanismo que creció en los últimos años, basado en la Web que permite escalar y virtualizar recursos de TI que son proporcionados como servicios a través de la red. Características inherentes y esenciales que deben ser provistas por las aplicaciones de cloud computing son: servicio bajo demanda, acceso ubicuo, escalabilidad, elasticidad, independiza al usuario del mantenimiento y pago por uso, siendo la seguridad todavía un desafío [1] [2]. Por otra parte, en los últimos años se ha logrado un progreso significativo en el uso de tecnologías de comunicación para almacenar y distribuir datos médicos bajo formatos digitales [3].

El uso de las mismas no es seguro cuando los datos, médicos o de otro origen, circulan libremente por redes abiertas como Internet, expuestos a que los mismos sean alterados o mal utilizados. Esto sucede especialmente cuando se emplean servicios de teleconsulta o telediagnóstico, que se están difundiendo por todo el mundo con el aporte de las Tecnologías de Informática y Comunicaciones. Un esquema de implementación se muestra en la Figura 1.

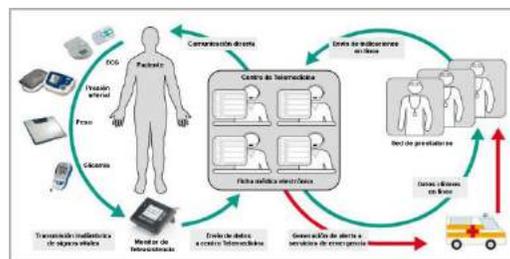


Figura 1. Típica red de telemedicina

La necesidad de tomar medidas de seguridad se incrementó con el almacenamiento de información médica en la nube.

Almacenar en la nube es una buena alternativa ya que permite a los centros médicos desentenderse del hardware y software usado, pero conlleva un mayor riesgo de violación de autenticidad e integridad en los registros del paciente. La confidencialidad de los datos es una necesidad ética en el campo de la salud. La tecnología de encriptación clásica es una herramienta importante que puede y debe ser utilizada para proteger los datos transmitidos en redes de computadoras [4], pero no es suficiente para solucionar todos los problemas de protección de datos digitales.

Además se debe considerar que, en la actualidad, los sistemas de información no son centralizados, sino distribuidos y que, por lo tanto, el control también debe estar distribuido. La protección primitiva que controlaba el acceso, mediante claves, hoy no es suficiente.

El watermarking o marcado de productos multimedia, imágenes, videos, audio, gráficos, etc., se empezó a desarrollar desde la década del 90 como forma de protección de propiedad intelectual [5]. Consiste en embeber bits en el archivo, sea imagen, video o audio, de forma visible (audible) o invisible (no audible). Estos bits extra constituyen la marca y en las primeras implementaciones permitían identificar al propietario, utilizándose posteriormente para alcanzar otros propósitos como detección de adulteraciones, aseguramiento de integridad e incorporación de metadatos. Esta técnica se utilizó posteriormente para conseguir otros propósitos como detección de adulteraciones, aseguramiento de integridad e incorporación de metadatos. Así, en estos días, la marca de agua aparece como un medio eficiente para asegurar integridad y verificar autenticidad. Actualmente, se está adaptando esta técnica para imágenes médicas, siendo de

particular interés el embebido de los datos médicos del paciente en sus imágenes personales [6]. Resultados experimentales, mostraron que la aplicación técnica de marcas de agua a este tipo de imágenes las tornaron más resistentes a varios tipos de ataques [7]. Las amenazas, como la destrucción de sistemas de software y violación en los accesos, están surgiendo con frecuencia en la plataforma de la nube, por lo que se hace absolutamente necesario tomar medidas para contrarrestarlas [8] [9].

Las imágenes médicas son almacenadas por los siguientes dos propósitos:

- Diagnóstico
- Base de Datos (almacenamiento a largo plazo)

Las imágenes deben ser guardadas perfectamente sin ninguna pérdida de información antes de que el médico haga su diagnóstico. Deben ser almacenadas sin compresión o comprimidas mediante un algoritmo que no pierda información.

Al comienzo de los años 80 apareció el PACS (Picture Archiving and Communication System/Sistema de Comunicación y Almacenamiento de Imágenes). Inicialmente, los PACS se desarrollaron para cubrir necesidades específicas, tales como adquisición de datos y visión de estos en estaciones de trabajo de poca capacidad [10]. Rápidamente se extendieron en las instituciones médicas, pero presentaban problemas de interoperabilidad ya que eran desarrollados en forma independiente por distintos proveedores que seguían sus propias reglas. Surgió entonces la Norma DICOM (Digital Imaging and Communications in Medicine), desarrollada por el ACR (American College of Radiology) en conjunto con NEMA (National Electrical Manufacturers Association) [11]. Tras varios intentos se aprobó en 1993 y sufre una actualización constante. Esta norma define el acceso a la web, la estructura de intercambio, las capas de comunicación de datos y los comandos para manejo de imágenes médicas que deben respetar todos los fabricantes para obtener interoperabilidad. Su aceptación hizo que se “dicomizaran” los PACs. Actualmente los fabricantes de equipos para imágenes médicas siguiendo indicaciones de la norma los acompañan

de un CS (Conformance Statement) que asegura que cumplen con la misma. Esta norma permite el acceso remoto a archivos en formato DICOM (extensión dcm) utilizando los ya clásicos protocolos TCP/IP, emplean el Protocolo HTTP (Hypertext Transfer Protocol) o HTTPS (Hypertext Transfer Protocol Secure). Si bien aseguró interoperabilidad entre los distintos sistemas y demostró cierta flexibilidad en entornos que manejan imágenes médicas, no hizo un aporte significativo a la seguridad ni al acceso de datos por fuera de instituciones médicas

En 1996, se dictó en EEUU, la HIPAA (Health Insurance Portability and Accountability Act) que indica qué requisitos se deben cumplir para las transacciones de datos de salud con el objetivo de que los datos médicos se almacenen y se puedan recuperar a largo plazo, evitando abusos y fraude. Microsoft en 2007 y Google en 2008 ofrecieron portales Health a los usuarios que querían que sus EHRs estuvieran disponibles para sus servicios de salud y para ellos mismos. En 2010, IBM y Aetna en conjunto anunciaron un nuevo uso de la plataforma de cloud computing de IBM diseñada para ayudar a los profesionales de la salud a acceder rápidamente a la información del paciente: registros médicos, recetas, y datos de laboratorio recolectados de múltiples fuentes para crear un registro detallado del mismo. Se estima que en el año 2020 el 80% de los datos se habrá mudado a la nube. El uso de estas plataformas y otras permite a los centros médicos desentenderse de problemas técnicos (actualización y mantenimiento de software y hardware), económicos y legales relacionados al manejo de datos lo que le conviene más allá de los riesgos que corre. Entre los inconvenientes se encuentra su latencia, la dificultad para tener el servicio disponible todo el tiempo, y la seguridad [12]. Se debe tener especialmente en cuenta que los datos almacenados en la nube son vulnerables a ataques internos. La identidad y ubicación de intermediarios y de los proveedores de servicio está disimulada, oculta, por la nube.

La importancia del watermarking en imágenes médicas ya se destacaba dos décadas atrás [13], y sigue considerándose en la actualidad [14]. Es de particular interés que se embeban los datos del

paciente en sus imágenes médicas personales. Las amenazas, como la destrucción de sistemas de software y violación en los accesos, están surgiendo con frecuencia en la plataforma de la nube, por lo que se hace absolutamente necesario tomar medidas para contrarrestarlas.

Se recomienda en telemedicina, la combinación de watermarking con técnicas de criptografía clásicas [13-17]. La encriptación puede impedir problemas en los nodos intermedios, pero no en los puntos finales que deben poder descifrar los datos y si el proveedor del servicio confía, a su vez, en otros proveedores entonces los datos del usuario pueden ser leídos por muchas entidades en la nube. Lo que se precisa para incrementar la confianza en la nube siguiendo esta línea de razonamiento es algún mecanismo que pueda detectar y castigar cualquier problema relativo a la confidencialidad. El usuario final debe confiar en la entidad que administra los EHR (sea Azure, Amazon, Google Health, Microsoft Health Vault u otro proveedor de nube). El administrador de EHR debe tener los medios para detectar y castigar las violaciones a la confidencialidad que se hubieran producido. Las nuevas técnicas deben tratar especialmente de mitigar los riesgos de que la información sufra ataques internos en la nube [18].

La propuesta actual de investigación consiste en realizar una comparación entre diversas plataformas de cloud computing para alojar imágenes médicas marcadas mediante un mecanismo que combine watermarking y encriptación para lograr seguridad, integridad y autenticidad de datos médicos. Se propone utilizar un sistema de archivos distribuidos Hadoop, usar MapReduce, modelo de programación para el manejo de grandes bases de datos [19].

IV. OBJETIVOS Y LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto se inscribe dentro de los lineamientos de investigación en Seguridad Informática.

Para el desarrollo de este trabajo de investigación se aplicará el método empírico-analítico, que se basa en la experimentación y en la lógica empírica, junto a la observación de plataformas de Nubes para alojar imágenes médicas con marcas de agua.

El objetivo de este proyecto de investigación es analizar el estado del arte de cloud computing para servicio de almacenamiento de imágenes médicas con marcas de agua y difundir los resultados obtenidos para realimentar el proceso de desarrollo de los algoritmos de watermarking.

V. MATERIALES Y MÉTODOS

En nuestra investigación proponemos hacer frente al problema de la seguridad de los datos contenidos en imágenes médicas alojadas en Cloud Computing, utilizando la técnica de marca de agua en el EHR. Luego se envía la imagen con marca de agua al proveedor de la Nube. En particular se trabajará con imágenes médicas en formato dicom y con marcas reversibles indetectables.

Es común que una imagen médica sea diagnosticada antes de que la misma sea almacenada en un almacenamiento a largo plazo, de este modo la parte significativa de la imagen, conocida como ROI (Region of Interest), es determinada en ese momento, con lo que el embebido de información extra se puede hacer fuera de esta zona.

VI. FORMACIÓN DE RECURSOS HUMANOS

El grupo está compuesto por una Directora, cuatro profesores investigadores de apoyo, tres ingenieros aspirantes a incorporarse a la carrera de investigador y un estudiante investigador de la carrera de Ingeniería en Sistemas de Información. Este proyecto contribuirá a la formación y crecimiento de la carrera de investigador de los integrantes del mismo.

Además existe la colaboración de una docente investigadora de la FCEFyN-UNC.

El desarrolla tareas de investigación en el Laboratorio de Investigación de Software (Lis).

Se dirigirán trabajos finales sobre la temática abiertos a estudiantes de Ingeniería en Sistemas de Información.

REFERENCIAS

- [1] Youssef *et al.* "Toward a Unified Ontology of Cloud Computing". Grid Computing Environment Workshop, IEEE, 2008.
- [2] Jadeja and Modi. "Cloud Computing – Concept, Architecture and Challenges".

- International Conference on Computing, Electronics and Electrical Technologies, IEEE, 2012.
- [4] Stallings. "Cryptography and Network Security", Ed. Prentice Hall, 4th Ed., 2005.
- [5] Cox, Miller y Bloom - "Digital Watermarking" - Morgan Kaufmann, 2002.
- [6] Coatrieux et al- "Relevance of Watermarking in Medical Imaging"- Proceedings of the IEEE EMBS Conf. on Information Technology Applications in Biomedicine, Arlington, USA, Nov. 2000, pp 250-255.
- [7] Fatma E.-Z. A. Elgamal, Noha A. Hikal, F.E.Z. Abou-Chadi- "Secure Medical Images Sharing over Cloud Computing environment", Information Technology dept. Faculty of Computers and Information Sciences, Mansoura University Mansoura, Egypt . 2013.
- [8] Junning Fu et al- "A Watermark-aware Trusted Running Environment for Software Clouds", School of Software Tsinghua University Beijing 100084, China, 2009.
- [9] Yu-Chao Liu et al- "A Method for Trust Management in Cloud Computing: Data Coloring by Cloud Watermarking", Department of Computer Science and Technology, Tsinghua University, Beijing, 2011.
- [10] Bharath. "Introductory Medical Imaging". Ed. John Enderle, University of Connecticut, 2009.
- [11] Pianykh. "Digital Imaging and Communications in Medicine (DICOM). A Practical Introduction and Survival Guide". 1st Ed., Ed Springer, 2008.
- [12] Zhang and Liu. "Security Models and Requirements for Healthcare Application Clouds". 3rd International Conference on Cloud Computing", IEEE, 2010.
- [13] Coatrieux *et al.* "Relevance of Watermarking in Medical Imaging". Proceedings of the IEEE EMBS Conf. on Information Technology Applications in Biomedicine, Arlington, USA, pp 250-255, 2000.
- [14] Aminzou *et al.* "Towards a Secure Access to Patient Data in Cloud Computing Environments". Security Days (JNS3), IEEE, 2013.
- [15] Elgamal, Hikal & Abou-Chadi. "Secure Medical Images Sharing over Cloud Computing environment". International Journal of Advanced Computer Science and Applications (IJACSA), vol. 4, N°5, 2013.
- [16] Bouslimi and Coatrieux. "A Joint/Encryption Watermarking System for Verifying the Reliability of Medical Images". Medical Data Privacy Handbook pp. 493-526, Springer, 2015.
- [17] Al-Haj, Hussein and Abandah. "Combining Cryptography and Digital Watermarking for Secured Transmission of Medical Images". Second International Conference on Information Management (ICIM), IEEE; 2016.
- [18] Garkotti *et al.* "Detection of Insider attacks in Cloud based e-healthcare". International Conference on Information Technology, IEEE, 2014.
- [19] Lee *et al.* "Implementation of MapReduce-based Image Conversion Module in Cloud Computing Environment". Proc. of Int. Conference on Advances in Computing, 2011.

ANÁLISIS SISTEMÁTICO DE LA SEGURIDAD EN INTERNET OF THINGS

Mg. Ing. Norma Beatriz Perez⁽¹⁾, Miguel Alfredo Bustos^{(1)(*)}, Dr. Mario M. Berón⁽¹⁾ & PhD. Pedro Rangel Henriques⁽²⁾

⁽¹⁾Departamento de Informática - Facultad de Ciencias Físicas Matemáticas y Naturales (FCFMyN) - Universidad Nacional de San Luis

⁽²⁾Universidade do Minho - Braga, Portugal

{nbperez, mabustos, mberon}@unsl.edu.ar⁽¹⁾, pedrorangelhenriques@gmail.com⁽²⁾

(*) Estudiante de la Carrera de Ingeniería en Informática

RESUMEN

Internet of Things (IoT, siglas en inglés) es una innovación tecnológica con gran auge en la actualidad debido a la incorporación en múltiples sectores en la sociedad como en la salud (Smart Health), en la educación (Smart Education), en el hogar (Smart Home), en el transporte (Smart Transport), en la seguridad (Smart Security), en las ciudades (Smart City), entre otros sectores. IoT permite interconectar objetos integrados físicos (Smart Object) en diferentes redes de comunicación transmitiendo y recibiendo información. Sin embargo, debido a su evolución exponencial surge la problemática de la seguridad y privacidad que afecta directamente en el desarrollo y mantenimiento de la utilización sostenible de IoT.

En este trabajo se describe la línea de investigación que aborda el estudio y análisis de seguridad en IoT. Dicho estudio se fundamenta en el análisis de la seguridad en las diferentes capas de la arquitectura de la misma.

Palabras Claves: Internet of Things, Objetos Físicos, Seguridad en IoT, Arquitectura IoT.

CONTEXTO

La presente línea de investigación se enmarca en el Proyecto (PO/16/93) de *“Fortalecimiento de la Seguridad de los Sistemas de Software mediante el uso de Métodos, Técnicas y Herramientas de Ingeniería Reversa”*. Realizado en conjunto con la Universidade do Minho Braga, Portugal. Recientemente aprobado por el Ministerio de Ciencia Tecnología e Innovación Productiva (Mincyt). Y por el Proyecto (P-031516.) de Investigación: *“Ingeniería de Software: Conceptos, Prácticas y Herramientas para el Desarrollo de Software con Calidad”* de la Facultad de Ciencias Físico Matemáticas y Naturales, Universidad Nacional de San Luis. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los

cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el Programa de Incentivos.

1. INTRODUCCIÓN

Denning [1] definió la seguridad de los datos como la ciencia que estudia métodos de protección de datos en los sistemas de red, e incluye controles criptográficos, controles de acceso, controles de flujo de información, controles de inferencia y procedimientos para respaldo y seguridad. La seguridad en sistemas basados en IoT implica no sólo proteger datos, claves criptográficas y credenciales. Es por esto, que estos sistemas son propensos a un amplio abanico de amenazas y desafíos en cuanto a la seguridad.

En IoT, los objetos físicos como por ejemplo automóviles, televisores, aires acondicionados, etc. que nos rodean son identificables de forma única y están interconectados [2]. A través de la red de comunicación los objetos se conectan entre sí recolectando información útil entre ellos. La información es transmitida a los diferentes dispositivos que tomarán acción ejecutando una tarea.

Debido a la aceptación de la incorporación de estos dispositivos integrados interconectados en la sociedad, se espera que para el 2020 [3] existan alrededor de 25 mil millones de objetos interconectados en la red mundial. En este marco, surgen amenazas de seguridad y privacidad poniendo a todos esos dispositivos en alto riesgo de vulnerabilidad, haciendo que dichos dispositivos funcionen beneficiando a los atacantes y no para el fin que se espera sean utilizados. Por esto, esta línea de investigación se centra en detectar, encontrar y proponer soluciones que sean adecuadas para reducir las amenazas que estos dispositivos pueden sufrir debido a su fácil acceso y manipulación [4]. Analizar y garantizar la seguridad en los dispositivos es una tarea prioritaria ya que los mismos tienen un impacto directo en la vida de los usuarios pudiendo invadir su privacidad. Para llevar adelante esta tarea, se requiere analizar y

garantizar la existencia de una infraestructura de seguridad correctamente definida, protocolos que limiten las posibles amenazas relacionadas con la escalabilidad, disponibilidad y seguridad en IoT [5].

La línea de investigación comprende los principales lineamientos de:

- Seguridad: Se centra en el estudio de los ejes fundamentales de la seguridad en IoT.
- Arquitectura: Se estudia cada capa y las problemática de la seguridad que pueden ocurrir en dichas capas.
- Modelos de Seguridad: Se estudian diferentes modelos de seguridad que permiten contrarrestar las vulnerabilidades producidas en la diversidad de dispositivos en el mercado de IoT.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

2.1 Seguridad en IoT

Los principales objetivos de seguridad en IoT son garantizar mecanismos de autenticación de identidad adecuados y proporcionar confidencialidad sobre los datos. Un modelo conocido para desarrollar mecanismos de seguridad en IoT se basa en tres áreas:

- **Confidencialidad de Datos:** Es la capacidad de proporcionar confianza y tranquilidad al usuario sobre la privacidad y confidencialidad de sus datos los cuales deben estar protegidos en su totalidad. Para esto se utiliza los mecanismos como el de cifrado de datos, verificación de dos pasos, entre otros.
- **Integridad de Datos:** Se refiere a la protección de datos relevantes, de hackers o de interferencias externas que se pueden producir durante la transmisión y recepción de datos que impide ser manipulados a menos que se detecte la amenaza a tiempo [7]. El Checksum y Cyclic Redundancy Check (CRC), entre otros, son mecanismos que garantizan esta actividad.

- **Disponibilidad de Datos:** Garantiza el acceso inmediato a sus recursos de información a la parte autorizada no sólo en condiciones normales sino también en condiciones adversas. Métodos de copia de seguridad de redundancia y conmutación por error que proporcionan la duplicación de los componentes del sistema en condiciones de falla del sistema o diversos conflictos del sistema que garantizan la fiabilidad y disponibilidad de los datos.

2.2 Arquitectura en IoT

En la Figura N° 1 se observa la arquitectura general de IoT que está compuesta por cuatro capas.



Figura N°1 Arquitectura general de IoT

A continuación se describen las diferentes partes que componen la arquitectura de IoT.

- **Capa de Percepción (CP):** Realiza la identificación de objetos. Recompila datos a través de sus sensores [7]. La seguridad se ve afectada debido a los sensores empleados en esta capa de percepción que son, generalmente, de diferentes tecnologías, por ejemplo los sensores RFID los cuales están expuesto a amenazas de:

Etiquetas: La falta de un mecanismo de autenticación adecuado en la mayoría de los RFID, permite el acceso de etiquetas sin estar autorizados. Lo que permite realizar lecturas, modificaciones e eliminaciones de datos [8].

Clonación de Etiquetas: Se crea una réplica de la etiqueta y se compromete

de manera tal que lector no pueda distinguirla de la etiqueta original [9].

Spoofing: Se difunde información falsa en los RFID haciendo parecer que la fuente es original [10]. De esta forma el sistema se vuelve vulnerable.

- Capa de Red: Transmite datos obtenidos de la capa de percepción a través de Internet, red móvil o cualquier otro tipo de red de comunicación confiable [11]. Las problemáticas de en la capa se red son:

Ataque Sybil: A un único nodo se le da múltiples identidades. Esto hace que el sistema de cómo resultado información incorrecta [12].

Ataque de Privación de Sueño: Mantiene a los nodos encendidos agotando la batería lo que lleva a que los nodos se apaguen [13].

Inyección de Código: Se inyecta código malicioso en un nodo. Esto podría provocar tener el control de la red [14].

- Capa de Nivel Medio: Es la encargada de garantizar el mismo tipo de servicio entre los objetos físicos conectados [15]. Los problemas de seguridad se producen en el canal de comunicación. Problemas de:

Acceso no Autorizado: Podría ser fatal para el sistema ya que el atacante puede prohibir el acceso a servicios relacionados de IoT, o incluido eliminar datos sensibles del sistema.

Ataque DoS: Ocasiona el apagado del sistema haciendo que los servicios no se encuentren disponibles.

- Capa de Aplicación: Es la encargada de las aplicaciones de IoT provenientes de los más diversos tipos de industrias, como por ejemplo, Smart Hospital, Smart City, Smart Transportation, entre otros [16]. La seguridad de esta capa se encuentra afectada por:

Inyección de código malicioso: Inyectar código malicioso en el sistema para robar algún tipo de información.

Ataque de denegación de servicio (DoS): Intenta romper el sistema

defensivo y por lo tanto la privacidad de los datos del usuario, mientras engaña a la víctima haciéndole creer que el ataque real está sucediendo en otro lugar.

Ataque de Spear-Phishing: Suplanta el correo electrónico en el que la víctima, una persona de alto rango, es atraída a abrir el correo electrónico a través del cual el adversario obtiene acceso a las credenciales de esa víctima y luego, con un pretexto, recupera información más confidencial.

2.3 Modelos de Seguridad en IoT

La seguridad del software es la ciencia y el estudio de la protección de software (incluyendo datos de software) contra el acceso no autorizado [17, 18, 19]. Existe diversos principios, enfoques y técnicas para mejorar la seguridad del software de amenazas como, por ejemplo, ataques de overflow de buffer [20], ingeniería inversa y alteración [21], entre otros [22]. Existen modelos que permiten contrarrestar e identificar las amenazas:

- **Modelo de Amenaza de Red:** El atacante intenta obtener acceso privilegiado que le permita realizar acciones maliciosas [23, 24, 25, 26, 27] como, por ejemplo lanzar otros ataques, consumir recursos o recolectar información.
- **Modelo de amenaza interna:** El hacker cuenta con algún nivel de privilegios ya sea en la red o en hardware que ejecuta la aplicación de destino. Esto le permite obtener acceso a datos sensibles los cuales puede alterar y/o realizar un hurto de los mismos.
- **Modelo de amenaza de host no confiable:** El hacker es local y cuenta con todos los privilegios. El hacker puede eliminar restricciones que hubieran hecho para que que el sistema se encuentre en un estado estable, eliminar la protección contra copia, hurto de información relevante, manipular la licencia de datos o de la aplicación.

3. RESULTADOS OBTENIDOS/ESPERADOS

El trabajo de investigación de esta línea permitió obtener los siguientes resultados:

- A partir del estudio y análisis de los dispositivos digitales basados en la tecnología IoT se detectó diversas vulnerabilidades que pueden ser contrarrestadas utilizando métodos técnicas y herramientas centradas en la seguridad de los dispositivos.
- El estudio de la arquitectura de IoT permitió determinar las amenazas (perdida de datos, manipulación de datos, perdida de privacidad, etc.) producidas en cada capa. A fin de reducir estas inseguridades deben ser consideradas en los dispositivos IoT.
- Se determinó la existencia de diferentes modelos para detectar amenazas que permiten contrarrestar la seguridad de software.

Los investigadores de esta línea pretenden a través de investigaciones sustanciales sobre medidas de privacidad y seguridad, evoluciones de riesgo, arquitectura, problemas de autenticación, entre otros, proporcionar respuestas a problemas abiertos en este campo del Internet of Things antes de que sean implementados e incorporados a la sociedad. Así como desarrollar métodos, técnicas y herramientas que ayuden, en un mayor grado, al desarrollo estable de tecnologías seguras en este campo de investigación tan amplio del Internet of Things.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de profesionales de la UNSL que forman parte de la línea de investigación de este trabajo llevan adelante diferentes trabajos finales integradores de *Ingeniería en*

Informática, Ingeniería en Computación, Licenciatura en Ciencias de la Computación, y en un futuro próximo trabajos finales de especialización, tesis de maestría y doctorado. En particular, las investigaciones desarrolladas en este trabajo forman parte del lineamiento inicial de una tesis doctoral de uno de los autores para optar al grado de Doctor en Ingeniería en Informática en la UNSL.

5. BIBLIOGRAFÍA

- [1] D. Denning, *Cryptography and Data Security*, Addison Wesley, 1982
- [2] D. Singh, G. Tripathi, A.J. Jara, A survey of Internet of Things: Future Vision, Architecture, Challenges and Services, in *Internet of Things (WF-IoT)*, 2014
- [3] Gartner, Inc. It can be accessed at: <http://www.gartner.com/newsroom/id/2905717>
- [4] Rodrigo Roman, Pablo Najera and Javier Lopez, "Securing the Internet of Things," in *IEEE Computer*, Volume 44, Number 9, 2011, pp. 51-58
- [5] Friedemann Mattern and Christian Floerkemeier, "From the Internet of Computers to the Internet of Things," in *Lecture Notes In Computer Science (LNCS)*, Volume 6462, 2010, pp 242-259
- [6] Luigi Atzori, Antonio Iera, Giacomo Morabito, "The Internet of Things: A Survey," in *Computer Networks*, pp. 2787-2805
- [7] Ying Zhang, *Technology Framework of the Internet of Things and Its Application*, in *Electrical and Control Engineering (ICECE)*, pp. 4109-4112
- [8] Mr. Ravi Uttarkar and Prof. Raj Kulkarni, "Internet of Things: Architecture and Security," in *International Journal of Computer Application*, Volume 3, Issue 4, 2014
- [9] Mike Burmester and Breno de Medeiros, "RFID Security: Attacks, Countermeasures and Challenges."
- [10] Aikaterini Mitrokotsa, Melanie R. Rieback and Andrew S. Tanenbaum, "Classification of RFID Attacks."
- [11] Xue Yang, Zhihua Li, Zhenmin Geng, Haitao Zhang, A Multilayer Security Model for Internet of Things, in *Communications in Computer and*

- Information Science, 2012, Volume 312, pp 388-393.
- [12] John R. Douceur, "The Sybil Attack," in Peer-to-Peer Systems - IPTPS, 2002, pp. 251-260
- [13] Tapalina Bhattasali, Rituparna Chaki and Sugata Sanyal, "Sleep Deprivation Attack Detection in Wireless Sensor Network," in International Journal of Computer Applications, Volume 40, Number 15, 2012
- [14] Priyanka S. Fulare and Nikita Chavhan, "False Data Detection in Wireless Sensor Network with Secure Communication," in International Journal of Smart Sensors and AdHoc Networks (IJSSAN), Volume-1, Issue-1, 2011
- [15] Rafiullah Khan, Sarmad Ullah Khan, R. Zaheer, S. Khan, Future Internet: The Internet of Things Architecture, Possible Applications and Key Challenges, in 10th International Conference on Frontiers of Information Technology (FIT 2012), 2012, pp. 257-260
- [16] Shi Yan-rong, Hou Tao, Internet of Things key technologies and architectures research in information processing in Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE), 2013
- [17] J. Viega, G. McGraw, Building Secure Software, Addison Wesley, 2001
- [18] M. Howard, D.C. LeBlanc, Writing Secure Code, 2nd ed., Microsoft Press, 2002
- [19] P.C. van Oorschot, "Revisiting Software Protection", pp.1-13, Proc. of 6th International Information Security Conference (ISC 2003), Bristol, UK, October 2003, Springer-Verlag LNCS 2851 (2003)
- [20] J. Wilander, M. Kamkar, "A Comparison of Publicly Available Tools for Dynamic Buffer Overflow Prevention", pp.149-162, Proc. of NDSS'03 (Internet Society): Network and Distributed System Security Symp., Feb. 2003, San Diego.
- [21] J. Gosler, "Software Protection: Myth or Reality?", Advances in Cryptology - CRYPTO'85, Springer-Verlag LNCS 218 (1985), pp.140-157
- [22] L. D'Anna, B. Matt, A. Reisse, T. Van Vleck, S. Schwab, P. LeBlanc, "Self-Protecting Mobile Agents Obfuscation Report", Network Associates Labs Report #03-015, 30 June 2003
- [23] CERT Advisory CA-1996-26 Denial-of-Service Attack via ping ("Ping of Death"), <http://www.cert.org/advisories/CA-1996-26.html>, accessed 29 Dec. 2003
- [24] D. Bleichenbacher, "Chosen Ciphertext Attacks against Protocols Based on RSA Encryption Standard PKCS #1", Advances in Cryptology - CRYPTO'98
- [25] D.M. Kienzle, M.C. Elder, "Recent Worms: A Survey and Trends", pp.1-10
- [26] N. Weaver, V. Paxson, S. Staniford, R. Cunningham, "A Taxonomy of Computer Worms", pp.11-18
- [27] Symantec, "What is the difference between viruses, worms, and Trojans?", <http://service1.symantec.com/SUPPORT/nav.nsf/pfdocs/1999041209131106>, October 1 2003

Tecnología Informática Aplicada en Educación

Dispositivos móviles como instrumento facilitador del aprendizaje

Mg. Roberto Bertone¹, Mg. José Luis Filippi², Lic. Guillermo Lafuente³, Mg. Carlos Ballesteros⁴,
Lic. Gustavo Lafuente⁵, I.S. Daniel Perez⁶, I.S. Sofía Aguirre⁷, A.S. Alejandra Mansilla⁸

LIAU⁹ - Facultad de Ingeniería – UNLPam.

pbertone@ada.info.unlp.edu.ar¹

{filippij², lafuente³, balleste⁴, gustavo⁵, perezd⁶, aguirres⁷, mansilla⁸}@ing.unlpam.edu.ar

⁹Laboratorio de Investigación de Ambientes Ubicuos

Resumen

El objetivo de este proyecto es mejorar los diferentes procesos que se efectúan en las instituciones educativas mediante la aplicación de las tecnologías de la información y la comunicación. En particular, se enfocará en la exploración y explotación de las tecnologías de la computación ubicua [1], direccionando la labor investigativa a mLearning.

Con el fin de explorar las posibilidades que brinda el cómputo móvil en el contexto educativo, se efectuará el trabajo experimental en la Facultad de Ingeniería - UNLPam, contando con un escenario de aplicación real.

Como parte del trabajo de investigación se realizarán diferentes tareas:

- Indagar el potencial de la tecnología móvil en el ámbito educativo.
- Analizar diferentes herramientas tecnológicas móviles en ambientes educativos.
- Entrevistar a los diferentes actores en su ámbito laboral cotidiano.
- Identificar diferentes escenarios de trabajo que muestren aspectos móviles en las actividades de sus actores.
- Documentar el proceso de investigación.
- Presentar en diferentes congresos de carácter nacional e internacional los resultados alcanzados.
- Confeccionar aplicaciones móviles de interés educativo.

- Difundir y capacitar a instituciones interesadas en utilizar las aplicaciones desarrolladas.

El proyecto trata de determinar la utilidad real de los dispositivos móviles en el ámbito educativo con el objetivo de establecer una implementación a gran escala.

Palabras claves: Aprendizaje Móvil. Dispositivo Móvil. Aprendizaje Colaborativo.

Contexto

Tipo de Investigación: Aplicada

Campo de Aplicación Principal: 7 1802 Computación, 7 1803 Comunicaciones.

Campos de Aplicación posibles: 13 1040 Ciencia y Tecnología, 7 4399 Otras – Educación – Tecnología Aplicada a la Educación

Institución que Coordina el Proyecto: Facultad de Ingeniería, Universidad Nacional de La Pampa.

Introducción

En ambientes con necesidades de representación y manejo de conocimiento como el ámbito educativo existen múltiples necesidades: desarrollo de contenido bibliográfico digital, implementación de programas que permiten la comunicación a través de variados dispositivos tecnológicos móviles, almacenar y distribuir el material educativo en variados formatos y diferentes plataformas, atender permanentemente a los actores involucrados dando respuesta a las

necesidades de padres, alumnos, docentes, administrativos y entidades del medio.

Estas relaciones constituyen en sí conocimiento agregado de suma utilidad, rebrotan la inquietud de Vannevar Bush (considerado el precursor de la Hipermedia), escrita en “*As we may think*”, donde establecía las dificultades para almacenar y consultar eficientemente la gran cantidad de conocimiento acumulado. La clave para que el material pueda ser consultado eficientemente está en la posibilidad de contar con las relaciones apropiadas que conecten dicho material.

Con la aparición de las redes de computadoras y la conectividad que ofrecen los dispositivos móviles, se conforma un ambiente propicio para satisfacer las necesidades de interacción entre los diferentes actores. Mark Weiser [1] menciona que: “*vamos camino a ambientes ubicuos, ambientes poblados de numerosos sensores que gracias a la miniaturización de los dispositivos son invisibles al usuario y están en permanente rastreo de la actividad humana*”. Aquí radica el objetivo primario del proyecto: indagar diferentes herramientas tecnológicas móviles a través de las cuales se puedan ofrecer servicios que satisfagan las necesidades de los usuarios que transitan en el ámbito educativo.

Situación Actual del Problema

En la sociedad actual se pueden observar una gran cantidad de componentes electrónicos que rodean a las personas y que están integrados en la vida cotidiana, gracias a su reducido tamaño y a su capacidad de cómputo, por ejemplo, computadoras personales, portátiles, teléfonos móviles de última generación, iPod, etc. [2], que logran comunicarse entre sí gracias a la existencia de protocolos de comunicación inalámbrica; a través de redes de celulares, redes de área

local (LAN), redes de área extensa (WAN), y Bluetooth entre otras [3].

Los dispositivos móviles son cada vez más pequeños, más poderosos y a medida que aumenta la demanda del consumidor, más baratos, incorporando nuevas funcionalidades diseñadas para permitir a los usuarios trabajar y acceder a la información en cualquier situación.[4]

Ante la presencia de nuevos dispositivos de comunicación un nuevo escenario social y educativo se define acorde a los cambios en la nueva organización del conocimiento, que inciden directamente en los procesos mentales. Está cambiando la forma de conocer, de interactuar con los demás, el volumen de información, su accesibilidad, los nuevos medios de almacenamiento masivo de información, nuevas formas de comunicarse a través de diferentes redes entre personas que se encuentran distantes físicamente. Mark Weiser clasifica la era de la computación en tres momentos: en 1980 el uso de las grandes computadoras (mainframes), en 1990 las computadoras personales (PC), y a partir del 2000 la computación ubicua (UC).

En este sentido, la tercera generación o paradigma en la computación cuenta con múltiples aplicaciones y equipos de investigación que procuran su desarrollo [5]. El concepto de ubicuidad se refiere en general a la presencia de una entidad en todas partes; pero en la computación adquiere la característica de ser, además, invisible. Este paradigma pretende brindar sistemas de cómputo inteligentes que se adapten al usuario y cuyas interfaces permitan que éste realice un uso intuitivo del sistema. De allí que la meta de la computación ubicua, de integrar varias computadoras y dispositivos inteligentes al entorno físico busca habilitar los beneficios de éstas y de la información digitalizada en todo momento y en todas partes.

La influencia de la computación ubicua es en gran parte de carácter teórico, y dadas las condiciones de integración entre educación y sistemas de cómputo explicada antes, comparte créditos junto a otros conceptos como usabilidad, interactividad e hipertextualidad [6]. Una posible aplicación en educación a distancia sería que el docente presentara el conocimiento en varias formas que se diferencien en sus características técnicas y didácticas. Así el estudiante optará por una vía conveniente de aprehensión del conocimiento con el consecuente mejoramiento de la calidad en el aprendizaje. Didácticamente, esto implicaría el diseño de distintos materiales o herramientas teniendo en cuenta los resultados de los antecedentes educativos del alumno. De esta manera se podrán ofrecer cambios significativos en el aprendizaje al conocer los modos de enseñanza que prefiere el estudiante. Otro aspecto que tiene relación con el concepto de ubicuidad y dispositivos móviles se encuentra en la ciencia cognitiva. En este sentido las teorías de la percepción retomadas por Norman (1998), hablan de la existencia de una percepción periférica donde los objetos solamente son apreciables cuando el ser humano centra la atención en ellos. En los sitios educativos se utilizan lenguajes de programación que posibilitan la “invisibilidad” de los elementos no relevantes en las acciones que ejecuta el usuario en la interfaz. Algunos de estos lenguajes, como XML [7], permiten que el receptor de la información pueda acceder a ésta desde cualquier navegador. Todo esto sucede sin percatarse de los cambios en el diseño que son necesarios para presentar los datos. A modo de ejemplo y bajo el mismo concepto, los futuros buscadores de contenidos que sigan los lineamientos definidos por la Web Semántica [8] permitirán arribar a contenidos con mayor precisión para el usuario, que los

buscadores tradicionales actuales; actual desafío de la llamada Web 3.0.

En el diseño de materiales didácticos para la web serán también importantes otras características técnicas, como por ejemplo la diferencia de peso entre archivos, la facilidad de lectura de un formato, el ancho de banda de la red y sus posibilidades multimediales, etc. [9]. Este conjunto de opciones técnicas que permiten la ubicuidad de un sistema de cómputo y las posibilidades conceptuales de diseñar la información pueden considerarse diseño ubicuo. Trascendiendo teóricamente de los sistemas computacionales a la concepción de aplicaciones didácticas en sitios educativos, el diseño ubicuo brinda opciones para mejorar el desempeño de Internet y la Web como medios de comunicación y educación [10].

A partir de establecer una relación estrecha entre el diseño de interfaz y el diseño didáctico, bajo el enfoque de la computación ubicua, ambos ofrecerán al usuario o alumno el conocimiento que requiere sin realizar un gasto cognitivo importante en el medio que transmite [11]. Una integración recíproca entre la educación a distancia y el diseño ubicuo, además de potenciar el uso de Internet en procesos de enseñanza y de aprendizaje, contribuirá a la difusión y aplicación de las tendencias tecnológicas que facilitan su uso. En este contexto la escuela debe cambiar, añadiendo nuevas herramientas tecnológicas que se encuentran presentes en la vida cotidiana de las personas (por ejemplo los teléfonos inteligentes) e incorporar su potencial al servicio de los diferentes actores, generando así un espacio de acción profesional distinto al actual. El entorno educativo cambia [12,3], la educación se presenta como la formación de los educandos en competencias, destrezas, habilidades para desempeñarse en un nuevo espacio social, el digital [13]. Nuevos escenarios educativos,

los mismos actores con un nuevo rol, y la implantación de las tecnologías móviles es el desafío actual [14].

Línea de Investigación y Desarrollo

El plan de actividades corresponde al proceso de investigación aplicada, con objetivos de corto, mediano y largo plazo, y una duración prevista de cuatro años.

Primer Año:

- Entrevistar a los diferentes actores, directivos, docentes, alumnos, personal administrativo, en su ámbito laboral cotidiano.
- Censar la presencia de dispositivos móviles, el tipo de plataforma utilizada.
- Investigar los requerimientos básicos de TICs para incorporar al proyecto.
- Indagar trabajos de terceros sobre las aplicaciones móviles utilizadas en el ámbito educativo.
- Analizar diferentes herramientas tecnológicas orientadas al desarrollo de actividades educativas móviles (mLearning).
- Definir una metodología de desarrollo de trabajo a efectuar.

Segundo y Tercer Año:

- Estudiar la tecnología necesaria para el desarrollo de aplicaciones móviles.
- Definir prioridades según necesidades de inmediatez para el uso y desarrollo de aplicaciones móviles orientadas a la educación.
- Confeccionar nuevas aplicaciones que se ejecuten en dispositivos móviles y que sean de interés educativo.
- Probar las aplicaciones desarrolladas en el ámbito educativo adecuado y realizar las pruebas necesarias para la puesta a punto.

Cuarto Año:

- Implantar las aplicaciones en los dispositivos móviles que hayan alcanzado un funcionamiento aceptable.
- Registrar el trabajo desarrollado en sus diferentes etapas.
- Difundir los logros alcanzados a la comunidad, para su posterior distribución.
- Presentar en diferentes congresos de carácter nacional e internacional los resultados alcanzados.
- Difundir y capacitar a instituciones interesadas en aplicar la tecnología presentada.

Resultados Obtenidos/Esperados

El proyecto da inicio a su actividad de investigación y desarrollo durante el año 2015. Entre los logros alcanzados durante los dos primeros años se destaca el desarrollo de una aplicación móvil denominada Facultad de Ingeniería Móvil, con el objetivo de crear una comunicación dinámica con los diferentes actores que transitan el ámbito universitario. Durante el tercer año se efectúa el desarrollo de una aplicación móvil llamada ReadMe, que permite a los usuarios acceder a documentos y haciendo uso del sintetizador de voz convertir el texto en audio. [15]

El objetivo, hacer de la institución un ambiente educativo ágil, que cumpla con las características básicas del aprendizaje móvil: ubicuo, flexible, portable, inmediato, motivante, accesible, activo, con conectividad permanente [16].

Con ello en mente, y a partir de diferentes prioridades que se evaluarán, se decide implementar aplicaciones móviles de interés a nuestro ámbito educativo, dando inicio a las tareas enmarcadas para el cuarto año del proyecto.

Al concluir el proyecto se espera contribuir en la instalación, personalización y/o desarrollo de diversas aplicaciones móviles que se distribuirán en forma libre y gratuita entre los diferentes actores que conforman el ámbito educativo universitario.

Formación de Recursos Humanos

Director de Proyecto

Co-Director de Proyecto

6 Investigadores

Referencias

[1] Weiser M. (1991), *The Computer for the Twenty-First Century*.

[2] Vedar E, Evans W, Griswold, W, (2009) *Ubibot - Prototyping Infrastructure for Mobile Context-Aware Computing*. Ubicomp 2009.

[3] Carmona M., González S., Castro Ruiz, *Innovación Tecnológica en Comunicaciones Móviles Desarrollada Con Software Libre: Campus Ubicuo* Novática: Revista de la Asociación de Técnicos de Informática, ISSN 0211-2124, N°. 190, 2007

[4] de la Riva D. (2007) “Aplicaciones Web para celulares”, Tesis de Grado. Facultad de Informática. Universidad Nacional de La Plata.

[5] Weiser, M. (1993), *Ubiquitous Computing*, IEEE Computer "Hot Topics", October 1993, <http://www.ubiq.com/hypertext/weiser/UbiCompHotTopics.html>

[6] Nielsen J. (2000), “*Designing Web Usability: The Practice of Simplicity*”, New Riders Publishing, Indianapolis, ISBN 1-56205-810-X

[7] W3C, Extensible Markup Language (XML), <http://www.w3.org/XML/>

[8] W3C, W3C Semantic Web Activity, <http://www.w3.org/2001/sw/>

[9] Girouard J, Horn H, Solovey S, Zibgelbaum, (2008), *Reality Based*

Interaction, a framework for post-WINP Interfaces.

[10] Krumm J., (2009), *Creating a New Book of Tutorials: Ubiquitous Computing Fundamentals*, Microsoft Research Redmond, Ubicomp 2009.

[11] Bravo C., Redondo M., Ortega M., Bravo J., *Evolución de un Entorno Colaborativo de Enseñanza Basado en Escritorio hacia la Computación Ubicua*, (2002) COLINE'02 Investigación en Entornos de Interacción Colectiva, Workshop de Investigación sobre nuevos paradigmas de interacción en entornos colaborativos aplicados a la gestión y difusión del Patrimonio cultural. Granada, 11 y 12 de Noviembre del 2002.

[12] Weiser, M. (1998) *The future of Ubiquitous Computing on Campus*. Communications of ACM, 41-1, January 1998, 41-42.

[13] Malani R., Griswold W, Simon B, (2009) *Public Digital Note-Taking in Lectures*. Ubicomp 2009.

[14] Richards M, Woodthorpe J, (2009), *Introducing TU100 "My Digital Life": Ubiquitous computing in a distance learning environment*. Ubicomp 2009.

[15] Filippi José Luis, Perez Hector Daniel, Aguirre Sofía, Bertone Rodolfo. (2017). ReadMe. Complemento de Aprendizaje Móvil. XXIII CACIC. <http://sedici.unlp.edu.ar/handle/10915/63626>

[16] M2learn, (2015). “M2learn, Framework for the development of mobile context-aware learning applications”. <https://code.google.com/p/m2learn/>

Determinación de perfiles de rendimiento académico en la UNNE con Minería de Datos Educativo

Julio C. Acosta^{1,2}, David La Red Martínez¹ Carlos Primorac¹

¹Facultad de Ciencias Exactas y Naturales y Agrimensura
Universidad Nacional del Nordeste

9 de julio N° 1449, (3400) Corrientes, Argentina.

julioaforever@hotmail.com, lrmdavid@exa.unne.edu.ar, carlosprimorac@gmail.com

²Facultad de Ciencias Agrarias / Universidad Nacional del Nordeste
J. B. Cabral N° 2131, (3400) Corrientes, Argentina.

Resumen

En este trabajo se propone evaluar el rendimiento de los estudiantes mediante técnicas de Minería de Datos. La propuesta no se enfoca en analizar el perfil del estudiante solo a través de sus calificaciones, sino también, estudiar el desempeño académico en base a otras variables.

Para definir los perfiles de los estudiantes y determinar patrones que conduzcan al éxito o fracaso académico, implementaremos un modelo que relaciona las calificaciones de los estudiantes con otras variables, tales como factores socioeconómicos, demográficos, actitudinales, entre otros; en base a lo cual clasificaremos los diferentes perfiles de alumnos.

Describimos el modelo a implementar con el uso de Data Warehouse para determinar los perfiles de rendimiento académico en las asignaturas Álgebra de la carrera Licenciatura en Sistemas de Información (LSI) de la Facultad de Ciencias Exactas y Naturales y Agrimensura (FaCENA) de la Universidad Nacional del Nordeste (UNNE) y Matemática I de la carrera Ingeniería Agronómica (IA) de la Facultad de Ciencias Agrarias (FCA) de la UNNE (PI 16F002 acreditado por Res. N° 970/16 CS).

Esperamos contribuir a encontrar una respuesta al bajo rendimiento académico de los alumnos observado históricamente, problema éste que es el disparador de nuestra investigación. Los modelos predictivos que buscamos, permitirán tomar acciones tendientes a evitar el fracaso académico, detectando los alumnos con perfil de riesgo de fracaso académico de manera temprana, a poco del inicio del cursado de las asignaturas; lo que permitirá concentrar en ellos los esfuerzos de tutorías y apoyos especiales.

Palabras clave: rendimiento académico; almacenes de datos; minería de datos; modelos predictivos.

Introducción

En el mejoramiento de la calidad académica en la Universidad, no necesariamente debe enfocarse sólo en el sistema de enseñanza-aprendizaje, sino que debe atender otras variables, como por ejemplo, la sistematización de procesos de evaluación permanentes que permitan monitorear cuestiones ligadas a la calidad académica y retroalimentar la propuesta de mejora para la Universidad (Briand et al., 1999). Uno de los factores más críticos que debe evaluarse continuamente es el rendimiento académico. Se define al rendimiento académico como la productividad del sujeto, matizado por sus actividades, rasgos y la percepción más o menos correcta de los cometidos asignados (Maletic et al., 2002). Al evaluar el rendimiento académico se analizarán elementos que influyen en el desempeño como: los factores socioeconómicos, la amplitud de programas de estudio, las metodologías de enseñanza, los conocimientos previos del alumno (Marcus, 2003); por esto, no resulta adecuado evaluar el desempeño general de los alumnos a través de porcentajes de aprobación, notas obtenidas, etc., ya que este proceso de evaluación no brinda toda la información necesaria que pueda ser utilizada para detectar, y corregir problemas cognitivos, de comprensión, de discernimiento, actitudinales. Implementamos un mecanismo que nos permite determinar las características propias del estudiante analizando la existencia de patrones de comportamiento y de condiciones de los estudiantes que posibiliten la definición de los perfiles de alumnos. Actualmente existen varios métodos para determinar y clasificar patrones que se utilizan en el área de la Inteligencia Artificial y del Aprendizaje de Máquinas (del inglés Machine Learning - ML) (Marcus & Maletic, 2003). La Minería de Datos (del inglés Data Mining - DM), son procesos de descubrimiento de nuevas y significativas relaciones, patrones y tendencias en grandes volúmenes de datos utilizando técnicas de AI y

ML. Estas técnicas permiten extraer patrones y tendencias para describir y comprender mejor los datos y predecir comportamientos futuros. Un DW es una colección de datos orientada a un dominio, integrada, no volátil y variante en el tiempo para ayudar a tomar decisiones (Salton, 1989). Los DW surgieron por dos razones: a) la necesidad de proporcionar una fuente única de datos limpia y consistente para propósitos de apoyo para la toma de decisiones; b) la necesidad de hacerlo sin afectar a los sistemas operacionales (Molina López & García Herrero, 2006). En este trabajo se propone la utilización de técnicas de DM, con volúmenes no muy grandes de datos que oscilaran de cientos a miles, sobre información del desempeño de los alumnos de las cátedras Álgebra (LSI) FaCENA-UNNE y Matemática I de la FCA-UNNE.

Materiales y métodos

Trabajamos para detectar grupos de estudiantes en riesgo de fracaso en sus estudios, a fin de adoptar acciones proactivas frente al desgranamiento y el bajo rendimiento académico de los alumnos de primer año en la Universidad. La experiencia se realiza en las asignaturas Álgebra de la carrera LSI de la FaCENA de la UNNE y en Matemática I de la carrera IA de la FCA de la UNNE.

Si bien ambas asignaturas tienen régimen de acreditación similar, difieren en la carga horaria y los tiempos de dictado a saber: Álgebra (LSI) tiene 128 (ciento veintiocho) horas reloj de dictado de las cuales el 50% corresponde a teoría y el 50 % a trabajos prácticos en la modalidad cuatrimestral (corresponde al primer cuatrimestre de primer año de la carrera), mientras Matemática I (IA) tiene 96 (noventa y seis) horas reloj de dictado con idéntica distribución porcentual de tiempos de dictado de teoría y de trabajos prácticos, pero en la modalidad trimestral (corresponde al primer trimestre de primer año de la carrera)

En ambas asignaturas para alcanzar la condición de alumno regular, los alumnos deben asistir al menos al 75% de las clases de trabajos prácticos, que se dictan dos veces por semana en clases de 2 hs. cada una y deben aprobar 2 (dos) exámenes parciales cuyos contenidos son exclusivamente de trabajos prácticos; cada uno de ellos tiene su instancia de recuperación y para aquellos alumnos que hayan aprobado al menos 1 (uno) de los parciales en cualquiera de las 4 (cuatro) instancias disponibles, existe una instancia más para recuperar el examen que queda aún sin aprobar. Cualquiera de los exámenes parciales se aprueba con 60 (sesenta) puntos sobre 100 (cien) puntos posibles. La asistencia a clases de teoría es libre y se dictan dos veces por semana en clases de 2 hs. cada una.

Se acreditan las asignaturas con un examen final al que se accede en condición de alumno regular o de alumno libre; el alumno regular debe rendir en el examen final solamente los contenidos de teoría en un examen oral. El alumno que se presenta al examen final en condición de alumno libre, debe rendir un examen escrito de trabajos prácticos y tras aprobar esa instancia pasa al examen de teoría en condiciones similares a la antes mencionada.

Los porcentuales de los alumnos que regularizan las Álgebra y Matemática I no son los deseados; en el caso de Álgebra, de 320 alumnos inscriptos en los últimos 4 años, aproximadamente un 20% no alcanza a rendir el primer examen parcial en promedio y al final del cursado, regularizan la asignatura solo un 30% aproximadamente, en el caso de Matemática I el desgranamiento después del primer parcial no es tan evidente y el porcentual aproximado de alumnos regulares al final del cursado es del 40%.

La cantidad de alumnos que regularizan y/o que aprueban las asignaturas involucradas en este proyecto no es satisfactoria, consideramos que esa situación puede contribuir al desgranamiento y deserción de los alumnos en los primeros niveles de sus carreras. Es importante, por tanto, estudiar y determinar cuáles son las variables que inciden en el rendimiento académico a fin de poder establecer estrategias de acción pedagógicas que permitan mejorar dicho rendimiento.

Trabajaremos principalmente en el desarrollo de métodos que contribuyan a encontrar técnicas para la detección temprana de los alumnos que tendrán dificultades en sus estudios, a fin ofrecerles una contención y acompañamiento especial en el inicio de sus estudios Universitarios. Indagaremos aspectos tales como: a) diferencia del nivel de aprendizajes de contenidos previos en los alumnos, b) situaciones particulares personales de los propios alumnos, c) la capacidad de las cátedras para el seguimiento del aprendizaje de los alumnos, d) escasa motivación para el estudio de ciencias básicas y otros que puedan revelarse como incidentes en la problemática que nos ocupa y otros que serán detallados adelante.

Para recuperar contenidos en los grupos de riesgo detectados trabajaremos con materiales elaborados con nuevas tecnologías de la información (NTIC). Esto no debe desplazar ni sustituir las formas presenciales de enseñanza - aprendizaje, sino más bien situarnos en la posición de ofrecer alternativas diferentes para aquellos alumnos que requieren modelos diferentes para sus estudios y aprendizajes. Consideramos que las NTIC tienen el potencial para desempeñar un papel importante en la recuperación de contenidos al permitir un

abordaje más eficaz, en el sentido de permitirnos procesos de aprendizaje más profundos y más persistentes (Motschnig-Pitrik & Holzinger, 2002), mientras el peso de un aprendizaje efectivo permanece con las personas, sus capacidades y valores interpersonales (Derntl et al., 2011). En tal sentido, entendemos importante en nuestro trabajo el estudio que se efectuará en dos poblaciones aparentemente diferentes como son los alumnos de las carreras Licenciatura en Sistemas de Información y los alumnos de Ingeniería Agronómica, para determinar si los perfiles de los estudiantes varían según la elección de la carrera y medir las diferencias en la predisposición y adaptación para el trabajo y aprendizaje mediado con las NTICs (lo cual se confirmará o no).

En los últimos años se han realizado numerosos trabajos relacionados con la producción de contenidos; actualmente se tiene una concepción global e integral del e-learning (Nichols, 2003), en estos nuevos escenarios se incluyen la combinación del aprendizaje cara a cara y el soportado por medios tecnológicos (especialmente la Web), tal que las fortalezas de ambas configuraciones se puedan aprovechar y explotar. Este aprendizaje combinado (blended learning o b-learning) se considera de suma utilidad no sólo para las universidades sino también para la sociedad en general.

Nosotros, desde nuestros trabajos previos, hemos podido corroborar lo que oportunamente hemos formulado, que los docentes del siglo XXI deben incorporar definitivamente las NTICs como recursos didácticos, sin abandonar los tradicionales de tiza y pizarrón, pero deben conocer el uso de las NTICs con al menos en parte del potencial que ellas ofrecen (Acosta & La Red Martínez, 2012); algunas teorías psicológicas y pedagógicas consideran necesaria la inclusión del e-moderador o e-moderador, docente con habilidades especiales en las actividades online (Salmon, 2000); la actividad del docente tutor se transforma a veces en un hecho fundamental, ya que la manera en que se usa la tecnología puede transformarse en un factor de gran influencia en la calidad de la EA-EV (enseñanza - aprendizaje en entornos virtuales). Se debe trabajar entonces para lograr una forma de EA-EV que tome en cuenta las necesidades individuales, los intereses y estilos (Wenger et al., 2009).

En este proyecto de investigación, las variables que inciden en el rendimiento académico de los alumnos serán detectadas a fin de establecer, a través de los valores que ellas toman en cada caso, la población de alumnos en riesgo de fracaso, para establecer acciones tendientes a evitar el fracaso de cada uno de los alumnos, con las acciones que correspondan en cada caso

particular y/o de cada grupo detectado y disminuir así el posterior desgranamiento.

Data warehouse

Como soporte de los datos trabajaremos con Data Warehouse (DW); en informática, un almacén de datos (DW), es un sistema especial de bases de datos utilizado para el almacenamiento de datos y el procesamiento de los mismos para la presentación de informes y análisis de información, es considerado como un componente central de la inteligencia de organizaciones.

Un DW es un repositorio de datos que proporciona una visión global, común e integrada de los datos (Curto Días, 2010) (Figura 1) y presenta las siguientes características: a) Orientado a un tema: organiza una colección de información alrededor de un tema central. b) Integrado: incluye datos de múltiples orígenes y presenta consistencia de datos. c) Variable con el tiempo: se realizan fotos de los datos basadas en fechas o hechos. d) No volátil: sólo de lectura para los usuarios finales.

Detrás de la arquitectura de componentes del DW existe un conjunto de procesos básicos asociados: los ETL (del inglés Extract, Transform, Load – Extracción, Transformación y Carga). Los procesos ETL hacen referencia a la recuperación y transformación de los datos desde las fuentes orígenes cargándolos en el DW. En primer lugar los datos se analizan desde las fuentes y se extraen aquellos que serán de utilidad para el proceso en ejecución.

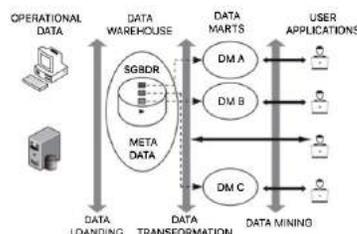


Figura 1 - Arquitectura Básica de un DW.

Luego de extraer los datos se los carga al DW pero, en muchas ocasiones, éstos requieren pasar por un proceso de transformación. La transformación de los datos significa un formateo y/o estandarización de los mismos convirtiendo ciertos números en fechas, eliminando campos nulos, etc.

Es necesario que antes de completar el DW con los datos se realicen controles para enviar información cualitativamente correcta. Luego se procede a aplicar alguna técnica para realizar el análisis de los datos almacenados en el DW. El método más utilizado es el proceso de DM que aplica la inteligencia artificial para encontrar patrones y relaciones dentro de los datos

permitiendo la creación de modelos, es decir, representaciones abstractas de la realidad.

Existen varias alternativas del DM, por ejemplo la Minería de Datos en Educación (Educational Data Mining, EDM). El objetivo de la EDM es el desarrollo de métodos para la exploración de tipos de datos únicos provenientes de plataformas educativas y usándolos para entender mejor a los estudiantes en el aprendizaje (Baker & Yaceff, 2009). Existen diversos estudios y publicaciones que abordan la evaluación de rendimiento académico utilizando técnicas de Minería de Datos (Formia & Lanzarini, 2013); (Pereira et al., 2013); (La Red Martínez et al. 2012); (La Red Martínez et al., 2017).

Modelo propuesto: La estructura del DW, se muestra en la Figura 2, consta de una tabla de hechos y varias tablas de dimensión. Una tabla de hechos o una entidad de hecho es una tabla o entidad que almacena medidas para medir el negocio como las ventas, el coste de las mercancías o las ganancias (IBM Knowledge Center, 2015).



Figura 2 - Modelo Propuesto del DW

Cada medida se corresponde con una intersección de valores de las dimensiones y generalmente se trata de cantidades numéricas, continuamente evaluadas y aditivas. Se pueden distinguir dos tipos de columnas en una tabla de hechos, columnas de hechos y columnas llaves. Las columnas de hechos almacenan las medidas del negocio que se quieren controlar y las columnas llaves forman parte de la clave de la tabla. Una tabla de dimensiones o entidad de dimensiones es una tabla o entidad que almacena detalles acerca de hechos. Por ejemplo una tabla de dimensión de hora almacena los distintos aspectos del tiempo como el año, trimestre, mes y día. Además incluye información descriptiva sobre los valores numéricos de una tabla de hechos. Las tablas de dimensiones para una aplicación de análisis de mercado, por ejemplo, pueden incluir el tipo de período de tiempo, región comercial y producto. Asimismo las tablas de dimensiones describen los distintos aspectos de un proceso de negocio. Si se desea determinar los objetivos de ventas, se pueden

almacenar los atributos de dichos objetivos en una tabla de dimensiones. Cada tabla de dimensiones contiene una clave simple y un conjunto de atributos que describen la dimensión.

En nuestro caso, las columnas de una tabla de dimensiones se utilizan para crear informes o para mostrar resultados de consultas. Por ejemplo las descripciones textuales de un informe se crean desde las etiquetas de las columnas de una tabla de dimensiones. El modelo que se presenta en este trabajo se compone de la tabla de hechos “ALUMNOS” y varias tablas de dimensiones asociadas a la misma que incluyen características que se desean estudiar. En la Figura 2 se representa gráficamente esta estructura.

Etapas de recolección de datos: Tal como se planteó, el estudio del desempeño académico de los estudiantes no sólo debe evaluarse teniendo en cuenta los resultados de las instancias de evaluaciones previstas por la asignatura sino que también deben analizarse otros factores culturales, sociales y/o económicos que afecten el rendimiento del alumno. Por ello para este trabajo resultó determinante la participación directa del estudiante, pues era necesario conocer datos sobre aspectos personales que no se podían obtener de otra manera que no fuera a través de respuestas directas por parte de cada alumno. A tal fin se dispuso la elaboración de una aplicación web que permitió contar con una Encuesta On-Line compuesta por preguntas relacionadas a situación familiar e historial de estudios secundarios, entre otras cuestiones.

Etapas de depuración y preparación de datos: Para la realización de una correcta explotación del DW se debe asegurar que los datos obtenidos en la etapa anterior sean consistentes y mantengan la coherencia entre ellos. Así, en la etapa siguiente, se realizará un proceso de limpieza en los datos, que es la eliminación de aquellos registros con todos sus campos en blanco, corrección de errores tipográficos, llenado de algunos campos nulos, entre otros. La Encuesta no permite la carga, por parte de los estudiantes, de calificaciones de la asignatura en estudio. Esto se dispuso así para evitar errores en los datos ya sea por olvido, o confusión al momento de ingresar los valores. Por ello la carga de notas correspondientes al primer parcial, segundo parcial y sus recuperatorios, examen final y situación del alumno (regular, promovido o libre), es realizada por el equipo responsable de este trabajo de investigación. La información se obtendrá a partir de la base de datos histórica de las cátedras con respecto a calificaciones de los alumnos. Con esta

información depurada se deberá proceder a trabajar en las próximas etapas: - Carga de Datos al DW: Mediante la ejecución del flujo de datos, la información almacenada en la tabla *encuesta* se distribuirá a las tablas pertenecientes al modelo del DW.

Resultados

Hasta el momento se ha completado la primera etapa que implicó el diseño del modelo del DW sobre el cual se implementarán técnicas de DM a fin de determinar perfiles de estudiantes vinculados a su desempeño académico en las asignaturas LSI-FaCENA e IA-FCA UNNE. En el avance que aquí se presenta respecto del Proyecto se pudo comprobar que la etapa de depuración y preparación de los datos ha demandado tiempo y esfuerzo debido principalmente a la poca integridad y coherencia que existía en la información que se utilizará para realizar la evaluación final. En etapas sucesivas se continuará con el proceso de minería de datos para evaluar y comparar patrones que se obtengan para definir los perfiles de estudiantes. La evaluación, análisis y utilidad de estos patrones con los que se construirá un modelo predictivo de rendimiento académico permitirá soportar la toma de decisiones eficaces por parte del cuerpo docente de las asignaturas involucradas

Referencias

Acosta, J. & La Red Martínez, D. (2012). Un aula virtual no convencional de Algebra en la FaCENA-UNNE. Saarbrücken: EAE.

Baker, R. & Yacef, K. (2009) The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, 1, 3-16.

Briand, L., Daly, J. & Wüst, J. (1999). A unified framework for coupling measurement in object-oriented systems. *IEEE Transactions on Software Engineering*, 25 (1), 91-121.

Curto Dias (2010). Introducción al business intelligence. UOC: Barcelona

Derntl, M., Hampel, T., Motschnig-Pitrik, R., & Pitner. (2011). Inclusive social tagging and its support in Web 2.0 services. *Computers in Human Behavior*, 27(4), 1460-1466.

Formia, S. & Lanzarini, L. (2013). Caracterización de la deserción universitaria en la UNRN utilizando minería de datos. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología (TE&ET)*, (11):92-98.

La Red Martínez, D.; Acosta J., Uribe V. & Rambo A.; (2012) Academic Performance: An Approach From Data Mining; V. 10 N° 1; *Journal of Systemics, Cybernetics and Informatics*; pp. 66-72; U.S.A.

La Red Martínez, D.; Giovannini, M.; Báez Molinas, M.; Torre, J. & Yaccuzzi, N. (2017) Academic performance problems: A predictive data mining-based model. *Academia Journal of Educational Research*; 5, 4 pp. 61-75. England, U.K.

Maletic, J., Collard, M. & Marcus, A (2002). Source Code Files as Structured Documents. *Proceedings 10th IEEE International Workshop on Program Comprehension (IWPC'02)*, 289-292.

Marcus, A. (2003). Semantic Driven Program Analysis, Kent, OH, USA, Kent State University Doctoral Thesis.

Marcus, A. & Maletic, J. (2003). Recovering Documentation-to-Source-Code Traceability Links using Latent Semantic Indexing. *Proceedings 25th IEEE/ACM International Conference on Software Engineering (ICSE'03)*. 3(10), 125-137.

Molina López, J. & García Herrero, J. (2006). *Técnicas de Análisis de Datos*. Madrid: Universidad Carlos III.

Motsching-Pitrik, R., & Holzinger, A. (2002). Student-centered teaching meets new media: concept and case study. *Journal of Educational Technology and Society*, 5(4), 160-172.

Nichols, M. (2003). A theory for e-Learning. *Journal of Educational Technology and Society*, (2), 1-10.

Salmon, G. (2000). E-moderating: The key to teaching and learning online. London: Kogan Page.

Salton, G (1989). *Automatic Text Processing: The Transformation, Analysis and Retrieval of Information by Computer*. Boston: Addison-Wesley Longman Publishing Co.

Pereira, R., Romero, A. & Toledo J. (2013) Descubrimiento de perfiles de deserción estudiantil con técnicas de minería de datos. *Vinculos*. (10) 1, 374-383 . Universidad distrital Francisco José de Caldas. Colombia.

Wenger, E., White, D. & Smith, J. D. (2009). Digital habitats. Stewarding technology for communities: Portland, OR, USA: Cpsquare.

Ambientes de Aprendizaje enriquecidos con Tecnología

Gustavo Alfredo Bacino¹, Lucrecia Ethel Moro, Stella Maris Massa, Adriana Laura Pirro y Hernán Hinojal

¹Facultad de Ingeniería/ Universidad Nacional de Mar del Plata/Argentina
(7600) Av. Juan B. Justo 4302, +54-223-481660
gustavobacino@gmail.com, lucreciamoro@gmail.com, smassa4@gmail.com,
adriana.pirro@gmail.com, hhinojal@fi.mdp.edu.ar

RESUMEN

Los ambientes de aprendizaje enriquecidos con Tecnologías de la Información y de la Comunicación (TIC) constituyen un contexto que exige un proceso de intervención pedagógica definido por los espacios, la organización social, las relaciones interactivas, una forma de distribuir el tiempo y un determinado uso de los recursos, donde los procesos educativos se desarrollan como elementos estrechamente integrados en dicho sistema desde la mediación tecnológica.

En este trabajo se describe el Proyecto de investigación “Ambientes de aprendizaje enriquecidos con Tecnología” que será desarrollado durante los años 2018/2019 en la Facultad de Ingeniería de la UNMDP.

El aporte de esta investigación está centrado en ofrecer un marco de interpretación de los aspectos de mayor incidencia, particularmente los procesos de interacción que se suscitan en estos nuevos escenarios educativos, ya que la inclusión de las TIC tiene cambios profundos en los modos de organizar, estructurar y consultar los contenidos educativos. Se plantean nuevos problemas y desafíos didácticos que podrán resolverse con nuevas estrategias para hacer frente a las exigencias cambiantes del entorno global y el énfasis del aprendizaje centrado en el estudiante a quien le toca asumir con mayor compromiso, de forma participativa y activa, su proceso de formación.

Palabras clave: Ambientes de aprendizaje, Tecnología, Competencias TIC.

CONTEXTO

Desde el año 2014 los integrantes del Grupo de Investigación en Tecnologías Interactivas (GTI) de la Facultad de Ingeniería de la Universidad Nacional de Mar del Plata vienen desarrollando tareas de investigación.

En el marco del Proyecto 15/G406 (2014/2015): Recursos Educativos Abiertos (REA) e Intervenciones de Gestión, Diseño e Implementación, se avanzó en la implementación de propuestas que involucran REA digitales en ambientes de aprendizaje. Se exploraron y construyeron instrumentos (rúbricas) para la evaluación de competencias en propuestas educativas con TIC [1], [2] y [3]. Se realizó una selección rigurosa de las producciones las cuales fueron compiladas y editadas en un libro digital: “Aprender con Tecnologías: estrategias de Abordaje” [4] que cubre un amplio espectro de estrategias que van desde la sistematización de los procesos de creación de objetos de aprendizaje considerando criterios de calidad pedagógicos, curriculares y tecnológicos; instrumentos de evaluación por competencias: la rúbrica o matriz de valoración y finalmente el diseño y la puesta en marcha de experiencias de enseñanza y aprendizaje con tecnología que potencian el desarrollo de competencias.

En el marco del Proyecto 15/G463 (2016-2017): Tecnología e Innovación en Ambientes de Aprendizaje: Desarrollo y Gestión, se avanzó en la construcción de ambientes de aprendizaje con videojuegos [5], [6] y en la exploración de las características de ambientes de aprendizaje de ciencias naturales enriquecidos con TIC [7], [8] y [9].

1. INTRODUCCIÓN

En las nuevas sociedades del conocimiento, donde la tecnología es la principal herramienta para gestionar la información, las personas tendrán la necesidad de formarse en ciencia y tecnología, para ello deberán adquirir las competencias para el manejo de las Tecnologías de la Información y la Comunicación (TIC); este tipo de habilidades o competencias son transversales para desempeñarse en cualquier área, tanto del mundo laboral como el de los estudios de nivel superior.

Los sistemas educativos están llamados a vivir cambios paradigmáticos en su actual configuración, y este proceso será facilitado y acelerado por el apoyo que presten las TIC para su desarrollo. El origen de un nuevo paradigma educativo es un esfuerzo por actualizar el sentido de la educación y las formas en que se desarrolla. Este paradigma se funda en la comprensión de todos los miembros de las comunidades educativas como aprendices. Ya no hay un conocimiento único y consolidado, transmitido desde los docentes, dueños del saber y del proceso de enseñanza, hacia estudiantes como receptores pasivos. Se trata ahora de una comunidad de personas que busca, selecciona, construye y comunica conocimiento colaborativamente en un tipo de experiencia que se conecta directamente con el concepto de comunidades de aprendizaje [10].

[11] afirma que es importante el lugar que ocupan hoy las nuevas tecnologías en relación con los modos en que se produce y difunde el conocimiento, y por ende es importante la necesidad epistemológica de su inclusión en las prácticas de la enseñanza.

En el contexto educativo, las TIC pueden ayudar a los estudiantes a adquirir las capacidades necesarias para llegar a ser competentes en su uso, buscadores, analizadores y evaluadores de información; solucionadores de problemas y tomadores de decisiones; ciudadanos informados, responsables y capaces de contribuir a la sociedad, entre otras, según la [12].

[13], [14] y [15] resaltan la potencialidad educativa de las TIC en tanto que estas

pueden posibilitar aprendizajes significativos y el desarrollo de las competencias antes mencionadas.

Al considerar las competencias TIC como fundamentales en cada saber, aparece la importancia de formar en estas habilidades. El docente es la persona que desempeña el papel más importante en la tarea de ayudar a los estudiantes a adquirir esas capacidades. Además, es el responsable de diseñar tanto oportunidades de aprendizaje como el entorno propicio en el aula que facilite el uso de las TIC por parte de los estudiantes para aprender y comunicar.

Distintas investigaciones destacan la importancia de generar ambientes de aprendizaje con TIC alrededor de problemas reales y de facilitar que los estudiantes trabajen en la solución de los mismos. Las TIC cumplen el papel de dar soporte a ambientes de aprendizaje que permitan al estudiante aprender y conectar sus aprendizajes con conocimientos previos o con otras disciplinas, experimentar, observar procesos y reflexionar acerca de ellos [16].

En este sentido, [17] sostiene que el apoyo que las tecnologías deben brindar al aprendizaje no es el de intentar la instrucción de los estudiantes, sino, más bien, el de servir de herramientas de construcción del conocimiento, para que los estudiantes aprendan con ellas, no de ellas. De esta manera, los estudiantes actúan como diseñadores, y los computadores operan como sus Herramientas de la Mente para interpretar y organizar su conocimiento personal. Las Herramientas de la Mente son aplicaciones de los computadores que, cuando son utilizadas por los estudiantes para representar lo que saben, necesariamente los involucran en pensamiento crítico acerca del contenido que están estudiando [18]. Las Herramientas de la Mente sirven de andamiaje a diferentes formas de razonamiento acerca del contenido. Es decir, exigen que los estudiantes piensen de maneras diferentes y significativas acerca de lo que saben. Desde esta perspectiva se asume que la inteligencia se encuentra distribuida, de forma que está situada no sólo en la mente de la persona, sino más bien

esparcida en diferentes elementos, medios y personas, es decir, en diferentes entornos simbólicos y en entornos físicos que rodean al sujeto. Siendo una de las habilidades y competencias que debe tener el estudiante, el trabajar cognitivamente con ellas y readaptarlas para resolver los problemas que se le vayan presentando.

Desde esta perspectiva, se presentan nuevos escenarios conjugados a las potencialidades que ofrecen las TIC, configurando así otros espacios para la acción educativa, en los que se pretende atender los procesos de enseñanza y aprendizaje individual, en comunidad, a través de la interactividad y la actividad colaborativa. Sin embargo, las discusiones y apreciaciones orientadas en este sentido, insisten en la necesidad de sistematizar y profundizar este tipo de experiencias, el impacto en cuanto a su utilización, los lineamientos para desarrollar pertinentemente esta práctica, los efectos sobre los procesos de enseñanza y aprendizaje, las interacciones comunicativas y las relaciones socio afectivas que se derivan entre los agentes culturales y las herramientas comunicativas en estos espacios pedagógicos.

La mirada que nos presenta [19] sobre las tecnologías educativas revela que hoy en día su ámbito de estudio son las relaciones e interacciones entre las TIC y la educación. Por tanto, asumir esta postura desde una racionalidad crítica y postmoderna del conocimiento, significará que cualquier análisis de los problemas educativos que tengan relación con lo tecnológico deberá ser interpretado desde posicionamientos no solo técnicos del conocimiento psicopedagógico, sino sobre el significado de la educación y de los procesos socioculturales de cambio.

2. LÍNEAS DE INVESTIGACIÓN

Las principales líneas de investigación que se abordan en este proyecto son:

- Interacción persona-ordenador
- Serious Games y Gamificación
- Entornos formativos potenciados por tecnologías interactivas

3. RESULTADOS ESPERADOS Y OBJETIVOS

El objetivo general del proyecto de investigación que se presenta en este artículo es: interpretar los procesos de interacción entre los actores y los elementos que conforman los ambientes de aprendizaje enriquecidos con TIC en los niveles educativos medio o universitario.

Para ello se han planteado los siguientes objetivos específicos:

- Explorar las características de los ambientes de aprendizaje enriquecidos con TIC.
- Analizar las prácticas de aprendizaje que se promueven en ambientes de aprendizaje con TIC.
- Identificar cuáles son las interacciones suscitadas entre los actores y los elementos que conforman los ambientes de aprendizaje enriquecidos con TIC.
- Describir los procesos de interacción generados entre los actores y los elementos que conforman los ambientes de aprendizaje enriquecidos con TIC.
- Comprender los procesos de interacción entre actores y los elementos que conforman los ambientes de aprendizaje enriquecidos con TIC.

Los aportes de esta investigación están centrados en ofrecer un marco de interpretación de los aspectos de mayor incidencia, particularmente los procesos de interacción que se suscitan en estos nuevos escenarios educativos, donde el acceso a nuevos soportes de la información y de las comunicaciones tiene cambios profundos en los modos de organizar, estructurar y consultar los contenidos educativos. En estos contextos/escenarios se plantean nuevos problemas y desafíos didácticos que podrán resolverse con nuevas estrategias para hacer frente a las exigencias cambiantes del entorno global y el énfasis del aprendizaje centrado en el estudiante a quien le toca asumir con mayor compromiso, de forma participativa y activa, su proceso de formación.

4. FORMACIÓN DE RECURSOS HUMANOS

La formación de recursos humanos se considera de vital importancia. Integrantes del proyecto se encuentran desarrollando y dirigiendo dos tesis de postgrado, de la UNR y de la UNMDP en el marco del proyecto de investigación, correspondientes al Doctorado en Humanidades y Artes- Mención Ciencias de la Educación y a la Especialización en Docencia Universitaria.

5. REFERENCIAS BIBLIOGRÁFICAS

[1] Bacino, G., Morcela A. y Moro, L. E. (2014a). “Análisis y validación de una rúbrica para la evaluación de la competencia “comunicación eficaz escrita” en asignaturas experimentales”, *Revista Argentina de Ingeniería RADI*, Vol IV, pp.91-96.

[2] Bacino, G., Morcela, A. y Moro, L. (2014b). “Análisis y validación de una rúbrica para la evaluación de la competencia “comunicación eficaz” escrita en asignaturas experimentales”, *II Congreso Argentino de Ingeniería (CADI 2014). VIII Congreso Argentino de Enseñanza de la Ingeniería (CAEDI 2014)*, San Miguel de Tucumán, Argentina, septiembre 2014.

[3] Pirro, A. y Fernández M. E. (2015). Mapas conceptuales y desarrollo de competencias. Límite funcional, En Massa, S. M., Moro, L. E. y Bacino, G. A. (Eds.), *Aprender con Tecnologías. Estrategias de Abordaje*. E-Book, 2015, DOI: 10.13140/RG.2.1.1073.9287, Universidad Nacional de Mar del Plata, Diciembre de 2015.

[4] Massa, S. M., Moro, L. E. y Bacino, G. A. (Eds.) (2015). *Aprender con Tecnologías. Estrategias de Abordaje*, 2015, E-Book, 2015, DOI:10.13140/RG.2.1.1073.9287. Universidad Nacional de Mar del Plata.

[5] Farías Y., Fornasier M. S., Moro L. y Morcela A. (2016). “Aprender energía a partir de un videojuego”. *3° Jornadas en Educación*. Mar del Plata, Argentina,

Universidad Nacional de Mar del Plata. Junio de 2016.

[6] Massa, S. M, Evans, F., Zapirain, E., Kühn, F., Hinojal, H., Fernández, M. E., Morcela, A., Moro, L. y Pirro, A. (2017). “Aprendizaje basado en videojuegos. Un proyecto de integración de la tecnología en educación.” *XI Congreso de Tecnología en Educación y Educación en Tecnología (TEYET 2017)*. Buenos Aires. Argentina.

[7] Moro L. y Massa S. M. (2016a). Aprendizaje de ciencias naturales mediado con TIC: estudio de caso de una experiencia innovadora. *XVII Encuentro Internacional Virtual Educa*, Puerto Rico.

[8] Moro L. y Massa S. M. (2016b). Exploración de las características de ambientes de aprendizaje de ciencias naturales enriquecidos con TIC. *3° Jornadas en Educación*. Junio de 2016, Mar del Plata, Argentina.

[9] Moro, L. y Massa, S. M. (2016c). El aprendizaje de ciencias en ambientes enriquecidos con tecnologías, un estudio interpretativo de las interacciones en la educación secundaria, *II Jornadas sobre Formación del Profesorado para el Nivel Superior: experiencias didácticas e investigación*, Mar del Plata, Argentina, 2015.

[10] UNESCO, (2013). *Enfoques estratégicos sobre las TICs en Educación en América Latina y El Caribe*, Santiago: OREALC/UNESCO.

[11] Maggio, M. (2012). *ENRIQUECER LA ENSEÑANZA. Los ambientes con alta disposición tecnológica como oportunidad*. Buenos Aires: Paidós.

[12] UNESCO (2008). *Estándares de competencias en TIC para docentes*. Publicado por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. Recuperado de <http://www.oei.es/tic/UNESCOEstandaresDocentes.pdf>.

[13] Cabero, J. (2007). *Nuevas tecnologías aplicadas a la educación*. Madrid: McGraw-Hill.

[14] Casamayor, G. (2008). La formación on-line una mirada integral sobre el e-learning, b-learning. España: Graó.

[15] Engel, A., Coll, C. y Bustos, A. (2010). Aprender y enseñar con las tecnologías de la información y comunicación en la educación secundaria. En C. Coll (Coord.) *Desarrollo, aprendizaje y enseñanza en la educación secundaria*. Barcelona: Graó.

[16] Jaramillo, P., Castellanos, S., Castañeda, C. P. y Ordóñez, C. (2006). Características de los ambientes de aprendizaje en el aula de informática. *VIII Congreso Colombiano de Informática Educativa – TIC en educación y su incidencia en el desarrollo social*. Cali (Valle). Recuperado de: <http://ribiecol.org/embebidas/congreso/2006/p/onencias/trabajos/25/PonenciaCongresoRIBIEmay2.pdf>.

[17] Jonassen (2004). Computadores como Herramientas de la Mente. *EDUTEKA*. Recuperado de: http://lets.cinvestav.mx/Portals/0/SiteDocs/MediatecaSS/lets_sur_mediateca_jonassen_computadoresherramientas.pdf

[18] Jonassen, D. H. y Reeves, T. C. (1996). Learning with technology: Using Computers as cognitive tools. In D. H. Jonassen (Ed), *Handbook of research for educational communications and technology* (pp. 693-719). New York: Macmillan.

[19] Área Moreira, M. (2009). *Introducción a la tecnología educativa*. España. Recuperado de <https://campusvirtual.ull.es/ocw/file.php/4/ebookte.pdf>.

Modelos y herramientas para el proceso de desarrollo de Serious Games

Stella Maris Massa¹, Felipe Evans², Adolfo Spinelli³, Zapirain, Esteban⁴, Carlos Rico⁵, Antonio Morcela⁶, Franco Kühn⁷

Facultad de Ingeniería/ Universidad Nacional de Mar del Plata/Argentina
(7600) Av. Juan B. Justo 4302, +54-223-481660
smassa@fi.mdp.edu.ar¹, evansfelipe@gmail.com², spinelliadolfo@gmail.com³,
estebanzapirain@fi.mdp.edu.ar⁴; crico@crico.com.ar⁵, omorcela@fi.mdp.edu.ar⁶,
fdkuhn@gmail.com⁷

RESUMEN

Se presenta un proyecto I+D+T, (Investigación, Desarrollo y Transferencia) bianual.

En este proyecto focalizamos nuestra investigación en los modelos y herramientas para el desarrollo de los Serious Games.

Un Serious Game es un videojuego donde el objetivo trasciende el mero entretenimiento y busca incorporar un nuevo conocimiento en el jugador. La efectividad radica en su nivel de inmersión visual, sonora y emocional con el usuario y en su interactividad constante.

A pesar de que los Serious Games pueden contribuir en la capacitación y entrenamiento en múltiples áreas, una barrera en el éxito de su adopción son las dificultades para medir los resultados del aprendizaje a través de su uso.

En particular, la integración de las Analíticas de aprendizaje en el diseño de Serious Games ofrece oportunidades para rastrear y analizar datos del comportamiento de los estudiantes sobre la base de su interacción individual o grupal e interpretar el proceso de aprendizaje. Contar con un modelo para el Proceso de Elicitación de SG contribuirá con la difusión de buenas prácticas en un sector en expansión y permitirá la apertura de líneas de investigación, en procura de técnicas y procesos de desarrollo en dominios donde el trabajo multidisciplinario es fundamental.

Palabras clave: Modelos, Proceso de desarrollo, Serious Game.

CONTEXTO

El proyecto “Modelos y herramientas para el proceso de desarrollo de Serious Games” está inserto en el Grupo de Investigación en Tecnologías Interactivas (GTI) de la Facultad de Ingeniería de la Universidad Nacional de Mar del Plata.

Corresponde a la convocatoria 2018-2019 realizada por la Secretaría de Ciencia y tecnología de la UNMDP que financia dicho proyecto.

1. INTRODUCCIÓN

El juego es una de las actividades inherente al ser humano y es de carácter universal. Trabajos como los de Connolly et al. [1] y de Urquidí [2] aportan fuertes evidencias tanto teóricas como empíricas, sobre el juego como medio para adquirir habilidades.

En particular, un videojuego es un software donde los jugadores participan en un conflicto artificial definido por reglas que transcurre dentro de una realidad simulada donde el jugador pone sus emociones en juego [3], [4]. Más allá de su definición, el videojuego es una de las actividades humanas con mayor

evolución y aceptación de los últimos años [5].

En particular, en este proyecto nos centramos en un caso particular de videojuegos, los Serious Game (SG).

Autores como Abt [6] y Sawyer y Smith [7], coinciden en definirlo como un videojuego donde el objetivo trasciende el mero entretenimiento y busca incorporar un nuevo conocimiento o modificar alguna conducta en el jugador. Para lo cual el desarrollo de un SG debe lograr combinar que los obstáculos sean el aprendizaje de una o más habilidades necesarias para la vida real y que a su vez mantenga el interés del jugador realizando la actividad como voluntaria.

Ben Sawyer es un protagonista clave en la redefinición y el posterior éxito de los SG: en primer lugar, en 2002, publica un artículo alentando el uso de SG para la mejora y comunicación de políticas públicas [7]; inmediatamente después, fue co-fundador de la "Serious Games Initiative", una asociación que tenía como fin promover el uso de juegos para propósitos serios; y en el año 2008, desarrolla la Taxonomía de SG, agrupando en un solo concepto a videojuegos y simuladores desarrollados con cualquier propósito y dirigido a diversas industrias [8].

La efectividad de los SG como herramientas de simulación y entrenamiento radica en su nivel de inmersión visual, sonora y emocional con el usuario, y en su característica de interactividad constante. La experiencia de juego permite a un usuario entrar en estado flow (de inmersión y concentración total) en un tiempo mucho menor que por otros medios [9].

Aunque las dificultades para medir los resultados del aprendizaje logrado a través del uso de las SG han sido una barrera en el éxito de la adopción de los SG en Educación [10],

[11],[12], [13] y capacitación o entrenamiento [14], [15].

En esa línea, Baalsrud Hauge, et.al [10] señalan que el alto rendimiento en un juego, sin embargo, no implica necesariamente un aprendizaje efectivo. En general, el juego está inherentemente vinculado con el rendimiento, que va con una actitud de lograr hitos y altas puntuaciones. En contraste, el aprendizaje a menudo requiere oportunidades de reflexión, repetición, pausas e incluso la preparación para cometer errores y aprender de ellos. Por lo tanto, en muchos aspectos el proceso de juego puede entrar en conflicto con el proceso de aprendizaje.

Este conflicto entre aprendizaje y desempeño será mayor a medida que los SG ofrezcan más opciones abiertas y libertad de movimiento a los estudiantes: como hoy en día muchos SG tienden a reflejar enfoques de resolución de problemas realistas y una amplia gama de habilidades relevantes para los profesionales de este nuevo milenio [16]. Aunque cabe destacar que la evaluación del aprendizaje en SG requiere de otras metodologías que produzcan evidencias válidas del aprendizaje basado en juegos.

Los intereses recientes en las denominadas "Learning Analytics" o Analíticas de aprendizaje (LA) pueden ser una solución para resolver esta situación. Fournier, Kop & Hanan [17] definen LA como la "medida, colección, análisis y presentación de datos sobre los estudiante o capacitandos y su contexto, con el propósito de comprender y optimizar el aprendizaje y los entornos en los que se produce".

Por lo tanto la integración de LA en el diseño de SG ofrece nuevas oportunidades para rastrear y analizar datos del comportamiento de los estudiantes sobre la base de su interacción individual o grupal, interpretar el proceso de aprendizaje, realizar

recomendaciones y personalizar el aprendizaje [18], [19].

En el análisis acerca del desarrollo de SG, Ampatzoglou y Stamelos [20] destacan que existen pocos trabajos referidos a sistematizar los procesos propios de la Ingeniería de Requerimientos para videojuegos. Esto se repite en la construcción de SG, pues no se evidencia la existencia de metodologías, directrices y mejores prácticas para desarrollar productos eficaces e integrados en las rutinas de aprendizaje y procesos formativos, así como la evaluación de su impacto [21].

Manrubia Pereira [22] divide el proceso productivo de un videojuego en tres fases: pre-producción, producción y post-producción. La Elicitación y Especificación del videojuego conforman la pre-producción. Se buscará que el producto de la Elicitación tenga una forma parecida a la que toman los videojuegos comerciales que es el documento llamado GDD (Game Design Document).

Este documento contiene información detallada sobre los objetos, reglas, entornos, contexto, estructura, narrativa, condiciones de victoria/derrota y la estética del juego [23]. La construcción del GDD es fruto de un trabajo interdisciplinario donde participan todos los stakeholders interesados en el diseño creativo del videojuego. El documento resultante es un entregable evolutivo que se actualizará a medida que el desarrollo avance [24].

El GDD de un SG será de apariencia y contenido similar a cualquier videojuego, sin embargo existen varios aspectos que se deben tenerse en cuenta. El contenido del GDD debe validarse respecto de los objetivos pedagógicos, el dominio de inmersión y las habilidades a adquirir.

Finalmente, contar con un modelo para el Proceso de Elicitación de SG contribuirá con la difusión de buenas prácticas en un sector en expansión y permitirá la apertura de líneas de

investigación, en procura de técnicas y procesos de desarrollo en dominios donde el trabajo multidisciplinario es fundamental.

2. LINEAS DE INVESTIGACIÓN

Las principales líneas de investigación que se abordan en este proyecto son:

- Metodologías y técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software
- Interacción persona-ordenador
- Serious Games y Gamificación

3. RESULTADOS ESPERADOS Y OBJETIVOS

El proyecto que se presenta en este artículo tiene como objetivo general: “Proponer un Proceso de Elicitación para Serious Games en donde se combine el contexto y el perfil de usuario para mejorar tanto la experiencia del juego como la educativa”.

En ese marco se establecieron los siguientes objetivos específicos:

- Determinar las tareas y secuencias que permitan conformar el proceso de Elicitación de requerimientos de los Serious Games.
- Elaborar una estrategia para la definición de requerimientos de los Serious Games.
- Identificar información relevante del proceso de aprendizaje de los estudiantes sobre la base de sus datos de interacción en un Serious Game.
- Establecer las técnicas de Validación de requerimientos.
- Desarrollar un trabajo de campo en el que se analizará la viabilidad y posibilidades del Proceso de

Elicitación de requerimientos para un Serious Game.

- Construir una solución de software que proporcione información sobre el aprendizaje de los estudiantes en un Serious Game para realizar Analíticas de Aprendizaje.

El diseño de un sistema de información implica, además de una Especificación estructural y funcional, una descripción precisa de la interacción de los usuarios finales con el sistema de Información.

Teniendo en cuenta el gran impacto que ejercen las interfaces de usuario sobre el éxito o fracaso de los productos software, esta línea de investigación ha combinado las experiencias en el modelado de la interacción de la comunidad de Ingeniería de Software y de la comunidad de Interacción Persona-Ordenador (IPO).

Cruz-Lara, Fernández Manjón y Vaz de Carvalho [25] señalan que hay una clara ausencia de una estrategia integradora que incluya los resultados, las organizaciones y personas que se han beneficiado de estas iniciativas para potenciar los conocimientos, la experiencia y este know-how avanzado.

Por lo tanto, este es el momento de sistematizar los enfoques en SG, que combinen la teoría, la investigación y la práctica.

Para ello, es necesario crear los medios para reunir una masa crítica que permita la creación de redes y comunidades y la creación de las condiciones para discutir y hacer recomendaciones, análisis y metodologías en SG.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo cuenta con varios integrantes que están realizando estudios de posgrado.

En la actualidad, dos integrantes están concluyendo su tesis de Maestría en Ingeniería de Software de la Facultad de Informática de la UNLP.

Otro integrante está cursando el Doctorado en Ingeniería Orientación Modelado y Simulación Computacional de la Facultad de Ingeniería de la UNMDP.

Contamos además con un becario de investigación.

Se dirigirán trabajos finales y prácticas profesionales supervisadas (PPS) de estudiantes de la Carrera Ingeniería en Informática.

5. REFERENCIAS BIBLIOGRÁFICAS

- [1] Connolly, T. M., Boyle, E. A., Mac Arthur E., Hainey T. y Boyle J. M. (2012). A systematic literature review of empirical evidence on computer games an serious games. *Journal Computers & Education*, 52(2), 661-686. Elsevier. Filadelfia. Pensilvania, EEUU
- [2] Urquidi, M.. & Tamarit Aznar, C. (2015). Juegos serios como instrumento facilitador del aprendizaje: evidencia empírica. *Revista Opción*, 31(3), 1201 - 1220. Universidad de Zulia, Venezuela.
- [3] Salen, K. y Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. The MIT Press.
- [4] Crawford, C. (2003). *Chris Crawford on game design* (pag. 31). New Riders, Indianapolis EEUU.
- [5] Theesa (2016), Essential facts about the computer and video game industry: 2016 sales, demographic and usage data <http://essentialfacts.theesa.com/Essential-Facts-2016.pdf>

- [6] Abt, C. (1970). *Serious games*. The Viking Press. New York, EEUU.
- [7] Sawyer, B., Rejeski, D. (2002). *Serious Games: Improving Public Policy through Game-based Learning and Simulation*.
- [8] Sawyer, B. y Smith, P. (2008). Serious games taxonomy. Serious Game Summit 2008. San Francisco, USA.
- [9] Hamari, J. y Koivisto, J. (2014). Measuring flow in gamification: Dispositional Flow Scale – 2. En *Computers in Human Behavior* (40).
- [10] Baalsrud Hauge, J., Berta, R., Fiucci, G., Fernández Manjón, B. (2014). Implications of learning analytics for serious game design. In *Proceedings of the 14th International Conference on Advanced Learning Technologies (ICALT)*, (p. 230-232). Athens, Greece.
- [11] Alvarez, J. & Michaud, L. (2008). *Serious Games – Advergaming, edugaming, training and more*. *IDATE Consulting and Research*.
- [12] Ulicsak, M. (2010). Games in Education: Serious Games-A Futurelab Literature Review. http://media.futurelab.org.uk/resources/documents/lit_reviews/Serious-Games_Review.pdf
- [13] de Freitas, S. & Liarokapis, F. (2011) *Serious Games: A New Paradigm for Education?*. M. Ma, et al. (Eds) *Serious Games and Edutainment Applications* (pp 9-23). UK: Springer.
- [14] Boinodiris, P. (2012). *Playing to Win, Serious games for Business*” In *The Bridge Linking Engineering and Society*, p. 35.
- [15] Freire, M., Serrano-Laguna, A., Manero, B., Martínez-Ortiz, I., Moreno- Ger, P., Fernández-Manjón, B. (2016). Game Learning Analytics: Learning Analytics for Serious Games. In *Learning, Design, and Technology* (pp. 1–29). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-17727-4_21-1.
- [16] Westera, W., Nadolski, R. & Hummel, H. (2014). *Serious Gaming Analytics: What Students’ Log Files Tell Us about Gaming and Learning*. *International Journal of Serious Games*, 1, 35-50.
- [17] Fournier, H., Kop, R., Hanan, S.. (2011). *The Value of Learning Analytics to Networked Learning on a Personal Learning Environment*. Publications Archive Canada.
- [18] Siemens, G. (2010). What are Learning Analytics?. <http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>
- [19] Horizon Report. Johnson, L., Adams Becker, S., Gago, D. Garcia, E., y Martín, S. (2013). *NMC Perspectivas Tecnológicas: Educación Superior en América Latina 2013-2018*. Un *IEEE VAEP RITA, 1* Análisis Regional del Informe Horizon del NMC. Austin, Texas: The New Media Consortium.<http://dx.doi.org/10.1016/j.entcom.2014.02.003>.
- [20] Ampatzoglou, A. y Stamelos I. (2010). Software engineering research for computer games: A systematic review. *Information and Software Technology*, 51(9), 888-901. Elsevier. Filadelfia Pensilvania, EEUU.
- [21] Catalano, C. E., Luccini, A. M. y Mortara, M. (2014). Best Practices for an Effective Design and Evaluation of serious games., *International Journal of serious game 1(1)*.
- [22] Manrubia Pereira, A. M. (2014): El proceso productivo del videojuego: fases de producción. *Historia y Comunicación Social*, 19, 791-805. Universidad Complutense de Madrid.
- [23] Scott, R. (2010). *Level UP - The Guide to Great Game Design*. John Wiley & Sons.
- [24] Bethke, E. (2003). *Game Development and Production*. Wordware Publishing.
- [25] Cruz-Lara, S., Fernández Manjón, B. y Vaz *IEEE VAEP RITA, 1* de Carvalho, C. (2013). Enfoques Innovadores en Juegos Serios. *IEEE VAEP RITA, 1* (1), pp.19-21.

PLATAFORMA GAMIFICADA QUE PERMITE LA RESOLUCIÓN DE DESAFIOS EN LEGUNAJE C EN EL ÁMBITO UNIVERSITARIO

Pablo Martín Vera, Edgardo Javier Moreno, Rocío Andrea Rodríguez,
Federico Ezequiel Valles, Santiago Tamashiro, Facundo Vogel

GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software)
Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Florencio Varela 1903, San Justo, Provincia de Buenos Aires
{pvera, emoreno, rocio.rodriguez, fvalles, stamashiro, fvogel} @unlam.edu.ar

RESUMEN

La gamificación consiste en utilizar las dinámicas, mecánicas y componentes propios de los juegos en espacios no lúdicos, con el objetivo de motivar a los usuarios para resolver una determinada tarea. En este trabajo se plantea la gamificación en el ámbito universitario, para motivar a los alumnos a realizar ejercicios vinculados con la resolución de problemas en el Lenguaje C en el marco de la materia de Elementos de Programación, que se ubica curricularmente en el primer año de las carreras de Ingeniería que ofrece la Universidad Nacional de La Matanza. La plataforma gamificada presenta los ejercicios de una forma distinta por medio de desafíos y permite la corrección automática de los mismos, tanto desde el punto de vista de la codificación como a través casos de prueba lógica, planificados para corroborar que la resolución cumple con lo requerido en la consigna. Lo que agrega la motivación y jugabilidad es justamente la gamificación por medio de elementos propios de los juegos, en este artículo se presentan los avances del proyecto junto con los nuevos elementos incorporados a la plataforma.

Palabras clave: Gamificación, Programación, Educación, Elementos de Juego, Lenguaje C

CONTEXTO

El presente proyecto de Investigación y Desarrollo corresponde al grupo de investigación GIDFIS (Grupo de Investigación, Desarrollo y Formación en

Innovación de Software) perteneciente al Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza. El grupo es interdisciplinar al estar formado por docentes de distintos departamentos y en él se encuentran formándose en actividades de I+D alumnos de la universidad de las carreras de Ingeniería.

1. INTRODUCCIÓN

El concepto de gamificación fue abordado por diversos autores que han trabajado en su definición y en precisar sus elementos [1], [2], [3]. La Gamificación consiste en utilizar las dinámicas, mecánicas y componentes propios de los juegos en espacios no lúdicos, lo que permitirá motivar a los usuarios para resolver una determina tarea. Puede ser aplicado el concepto de gamificación en diversas áreas, entre ellas para motivar a empleados de una empresa, incentivar a la comunidad a realizar determinadas acciones de bien común y por supuesto dentro del espacio áulico. En cuanto a la gamificación en educación existen diversos trabajos, pero mayormente aplicados a niveles inicial o medio, no así en el nivel superior. “Si bien la gamificación ya se está aplicando con éxito en algunas áreas tan diversas como el marketing, los recursos humanos, la gestión de relaciones con los clientes, o incluso a la formación de altos directivos, apenas se han planteado experiencias de trasladar lo positivo de las mecánicas de juego a la docencia Universitaria... Sin embargo, existe un

creciente interés de la comunidad científica en la aplicación de gamificación a muchas áreas de conocimiento, por lo que muchas de las experiencias existentes pueden tratar de ser exportadas al dominio de la docencia.” [4].

Cabe destacar que “...la gamificación no es convertir las asignaturas en un juego, ni simplemente poner una puntuación a cada tarea. El proceso de gamificar, al igual que el proceso de virtualizar una asignatura, supone un esfuerzo de planificación y seguimiento, diseño y realimentación constantes y adaptados a cada asignatura, grupo y curso” [5].

Este proyecto está basado en aplicar la gamificación para conseguir motivar a los alumnos, en el ámbito universitario, en el uso de una plataforma que permite la realización de ejercicios por fuera del tiempo de cursada. La resolución de los ejercicios no se plantea como obligatoria, allí la importancia de la gamificación. Motivar a los alumnos a realizar una tarea opcional para reforzar el aprendizaje de la materia, es un punto clave. Se proponen los ejercicios en forma de desafíos a superar, incluso es necesario lograr puntos de desafíos concretados para poder desbloquear otros, este ha sido un concepto que ha mejorado en gran medida la jugabilidad de la plataforma.

La plataforma es accesible desde web, permitiendo al alumno resolver los desafíos propuestos en el Lenguaje C, que es el que se utiliza en la materia Elementos de Programación en el primer año de las carreras de ingeniería. Dichos desafíos tienen una devolución sobre su concreción en forma automática, analizándose no sólo si hay errores de codificación sino también si estos cumplen con las consignas dadas (pruebas lógicas), una vez superado el desafío según los errores cometidos y solucionados junto con el tiempo de realización se asigna al alumno un puntaje. Lo que permite establecer rankings individuales de alumnos pero también por cursos. A su vez ir acumulando puntaje le permite al alumno poder

desbloquear desafíos especiales (los cuales tienen un candado, ver figura 1) y también a medida que se progresa en el sistema se consigue ganar insignias.

La plataforma está en uso desde el primer cuatrimestre del 2016 (su desarrollo puede profundizarse en [6]).



Figura 1. Pantalla de Desafíos Pendientes – Tercer Nivel

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Se investigará principalmente sobre los siguientes temas:

- Análisis de los elementos de juego a incorporar a la plataforma incluyendo componentes, mecánicas y dinámicas.

- Nivel de aceptación de los nuevos elementos y su impacto en la motivación de uso.

En cuanto al desarrollo se prevé:

- Mejorar la plataforma para el análisis de desafíos.
- Diseño y desarrollo de nuevos componentes en la plataforma web.

RESULTADOS OBTENIDOS/ESPERADOS

La plataforma se encuentra en uso, habiéndose publicado los resultados obtenidos en cuanto al uso de la misma [7]. Actualmente se está trabajando en la mejora de la plataforma, para lo cual se añadió:

- 1) Incorporación de Foto de Perfil: Esto permite darle una característica de personalización al sistema, mostrando la foto de los estudiantes en los rankings. Dicha foto debe ser validada por un

docente antes de ser visualizada por el resto de los estudiantes a fin de evitar contenido inapropiado.

- 2) Detección de errores en la codificación de la función e incluso utilización de elementos no permitidos en el nivel actual del ejercicio (ver figura 2);
- 3) Desafíos bloqueados: Se añadieron desafíos que inicialmente aparecen bloqueados hasta alcanzar un determinado puntaje (los cuales se muestran con un candado – ver figura 1 – y se produce una animación al desbloquearlos). De esta forma se premia al estudiante que obtuvo buenos puntajes porque va a acceder a más desafíos y además al que no le fue bien, se va a esforzar para juntar puntos en otros niveles para ir desbloqueando ejercicios.

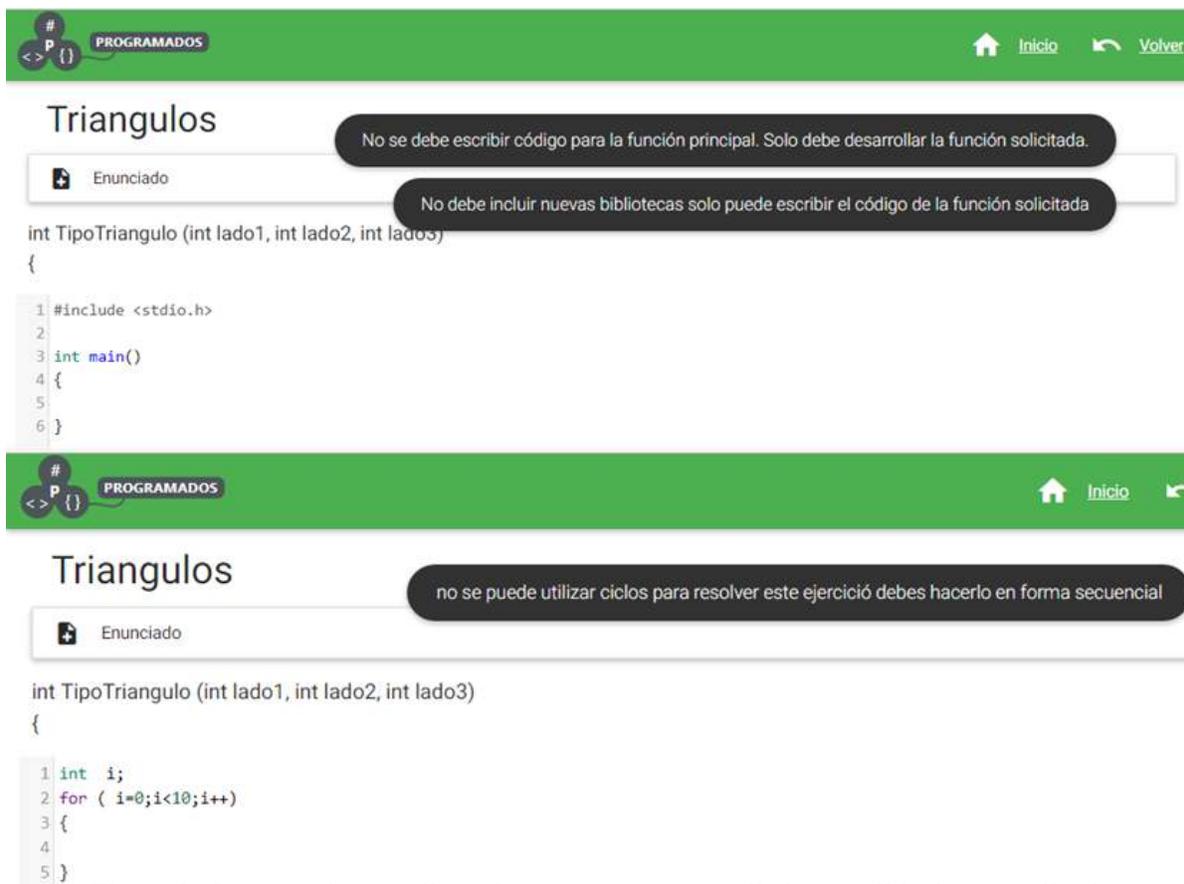


Figura 2. Capturas de pantalla en las que se presentan distintas validaciones realizadas

- 4) Insignias: Al finalizar todos los desafíos de un nivel se le dará al alumno una insignia. Para ello al finalizar cada ejercicio hay que verificar si con ese ejercicio finalizan todos los de dicho nivel y si es así se le otorga la insignia. La figura 9 muestra diseños posibles para las insignias de cada uno de los niveles.



Figura 3. Insignias incorporadas a la plataforma

Desde el menú principal se agregará un botón Logros. Debe llevar a una pantalla que muestre en gris las insignias pendientes de obtener y con color aquellas que ya ha ganado el alumno.

- 5) Mejoras gráficas y animaciones: se busca mejorar el aspecto general del sistema incorporando más elementos gráficos y animaciones que hagan a la plataforma más atractiva. Algunas de ellas son:
- Al desbloquear un desafío se muestra una pequeña animación de un candado que se abre.
 - Al ganar una insignia se muestra una animación con la insignia ganada.
 - En el listado de desafíos pendientes se agrega una visualización gráfica de los desafíos terminados y los pendientes mediante pequeñas medallas (ver figura 4).

Seleccionar Nivel

	Secuencial	0
	Decisión	0
	Ciclos	0
	Vectores	0
	Matrices	0
	Strings	2
	Estructuras	0
	Archivos	0
	Corte de Control	0

Figura 4. Listado de niveles con desafíos pendientes

- 6) Nivel de Usuario: Aún no está desarrollada esta funcionalidad pero se está considerando que cada alumno pueda ir alcanzando un determinado nivel de usuario en base a sus progresos en el sistema.
- 7) Tutorial de Uso: Si bien los docentes explican en el aula como ingresar al sistema y como se utiliza la herramienta, resulta necesario la realización de un tutorial de uso en donde pueda verse integrado a la plataforma como realizar un desafío. Así como muchos juegos incorporan en el primer nivel un tutorial en el que el usuario debe seguir indicaciones para hacer acciones con el sistema, esta plataforma se vería favorecida con un instructivo interactivo similar.
- 8) Vinculación con Redes Sociales: Dado que los alumnos utilizan con frecuencia diversas redes sociales, se planifica que los alumnos puedan configurar si desean o no que la plataforma publique las insignias obtenidas y otros logros alcanzados en sus redes sociales.

La plataforma permite brindar una nueva forma de ejercitar por fuera de la cursada. De este modo se consigue que los alumnos incrementen el tiempo que dedican a la materia y lo más valioso de la experiencia es que esto no se plantea como una actividad obligatoria, sino que es un recurso más, a disposición de los alumnos. Gamificar la plataforma ha sido parte del éxito dado que contribuye a la motivación de los alumnos, el interés por auto-superarse, la competencia entre cursos por alcanzar mejores puestos en el ranking general. Puede observarse comentarios sobre la herramienta en el recreo e incluso en alguna ocasión los alumnos le comentan al docente del curso que no pueden pasar a otro alumno. La implementación de los desafíos bloqueados también le añade un elemento que despierta un nuevo reto, conseguir los puntos necesarios para poder acceder a él. Hemos notado como el ranking general cambia constantemente y en algunos casos se puede ver como se intercalan posiciones, lo cual muestra una competencia constante. Las encuestas realizadas a fin de la cursada también demuestran el interés de los alumnos y el alto grado de aceptación de la herramienta. En los comentarios los alumnos agradecen la implementación de la plataforma e incluso proponen que esta iniciativa se promueva hacia otras materias. Se comenzarán con las tareas vinculadas al segundo año del proyecto con las cuales se conseguirá agregarle funcionalidades, mejorar e incorporar más elementos de gamificación, lo cual incidirá fuertemente en la jugabilidad de la misma e incrementar su uso.

3. FORMACIÓN DE RECURSOS HUMANOS

El grupo está formado por 10 personas:

- 5 docentes pertenecientes a diferentes departamentos conformando un grupo interdisciplinario.
- 1 asesora externa, en el área de gamificación perteneciente a la Universidad de La Laguna, España.
- 1 graduados de Ingeniería.

- 3 alumnos de ingeniería quienes se están formando en actividades de Investigación y Desarrollo.

Vinculado con esta temática se encuentra en realización una tesis de maestría.

4. BIBLIOGRAFIA

- [1] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining gamification," Proc. 15th ..., pp. 9–15, 2011.
- [2] Erenli K., "The impact of gamification: A recommendation of scenarios for education," in 2012 15th International Conference on Interactive Collaborative Learning (ICL), 2012, pp. 1–8.
- [3] Werbach, K (2013), "Gamificación". Fundació Factor Humà. Unidad de Conocimiento.
- [4] Cortizo Pérez, J. C., Carrero García, F. M., Monsalve Piqueras, B., Velasco Collado, A., Díaz del Dedo, L. I., & Pérez Martín, J. (2011). "Gamificación y Docencia: Lo que la Universidad tiene que aprender de los Videojuegos". http://abacus.universidadeuropea.es/bitstream/handle/11268/1750/46_Gamificacion.pdf?sequence=2&isAllowed=y
- [5] Vera, P. M., Moreno, E. J., Rodríguez, R. A., Vázquez, M. C., & Vallés, F. (2016, September). "Aplicación de técnicas de gamificación para la enseñanza de programación a alumnos de primer año de ingeniería". In XI Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2016).
- [6] Barragán Piña, A. J., Ceada Garrido, Y., Andújar Márquez, J. M., Irigoyen Gordo, E., Gómez Garay, V., & Artaza Fano, F. (2015). "Una propuesta para la motivación del alumnado de ingeniería mediante técnicas de gamificación".
- [7] Vera P., Moreno E, Rodríguez R, Vázquez M. Carina, Valles F, Cescon J (2017). "Gamificación en el Ámbito Universitario – Análisis e Implementación de Elementos de Juegos".

Diseño, modelado, simulación e implementación de técnicas emergentes de teoría de juegos en Serious Games

Esteban Aitor Zapirain¹, Stella Maris Massa¹

¹Facultad de Ingeniería/ Universidad Nacional de Mar del Plata/Argentina
(7600) Av. Juan B. Justo 4302, +54-223-481660
estebanzapirain@gmail.com, smassa4@gmail.com

Resumen

La teoría de juegos es una disciplina con múltiples aplicaciones en la toma de decisiones en el ámbito social, económico y político. La teoría cuántica de juegos se basa en los principios de la computación cuántica, añadiendo más flexibilidad y posibilidades a la teoría de juegos clásica.

Los serious games son juegos con un propósito educativo más allá del mero entretenimiento.

La teoría de juegos es de relevancia en las mecánicas de los videojuegos, en los cuales se desarrolla una competencia entre los factores del videojuego y el/los jugador/es humano/s.

El presente trabajo consta en la aplicación de los principios de la teoría cuántica de juegos en serious games. Las disciplinas en las que se basa este trabajo se consideran emergentes, por lo cual se espera que los resultados provoquen un alto impacto en el desarrollo de estas áreas del conocimiento.

Palabras clave: Serious games, Simulación, Modelo, Teoría de juegos, Teoría cuántica.

Contexto

Este artículo presenta la investigación realizada en el marco de la tesis: “Diseño, modelado, simulación e implementación de técnicas emergentes de teoría de juegos en serious games” del Doctorado en Ingeniería Orientación Modelado y Simulación

Computacional de la Facultad de Ingeniería de la Universidad Nacional de Mar del Plata.

Se inserta además en el proyecto: “Modelos y Herramientas para el Proceso de Desarrollo de Serious Games” (2018/2019) del Grupo de Investigación en Tecnologías Interactivas (GTI) de la citada Facultad.

Introducción

Serious games

El término juego alude a una actividad u ocupación voluntaria, ejercida dentro de ciertos y determinados límites de tiempo y espacio, que sigue reglas libremente aceptadas, pero absolutamente obligatorias, que tiene un final y que va acompañado de un sentimiento de tensión y de alegría, así como de una consciencia sobre la diferencia con la vida cotidiana [1].

El término serious games fue introducido por Abt en 1970 en su libro homónimo [2]. Su definición fue que estos juegos poseen un propósito educativo explícito y cuidadosamente planeado, y no están pensados para ser jugados principalmente por diversión.

En su época Abt se refería principalmente a los juegos de mesa y a los juegos de cartas, pero su definición puede extrapolarse a los juegos digitales. En 2005, Zyda [3] actualiza el término para referirse a videojuegos. El autor define al serious game como un desafío

mental, jugado con una computadora de acuerdo a reglas específicas, que usa el entretenimiento para promover la formación gubernamental o corporativa, con objetivos en educación, sanidad, políticas públicas y comunicación estratégica .

Ben Sawyer es clave en la redefinición y el posterior éxito de los serious games: en primer lugar, en 2002, publica un artículo alentando el uso de serious games para la mejora y comunicación de políticas públicas [4].

Inmediatamente después, fue co-fundador de la “Serious Games Initiative”, una asociación que tenía como fin promover el uso de juegos para propósitos serios; y en el año 2008, desarrolla la Taxonomía de serious games, agrupando en un solo concepto a videojuegos y simuladores desarrollados con cualquier propósito y dirigido a diversas industrias [5].

La efectividad de los serious games como herramientas de simulación y entrenamiento radica en su nivel de inmersión visual, sonora y emocional con el usuario, y en su característica de interactividad constante. La experiencia de juego permite a un usuario entrar en estado flow (de inmersión y concentración total) en un tiempo mucho menor que por otros medios [6].

Teoría de Juegos

La dificultad creciente en la secuencia de retos implica la existencia de una curva de aprendizaje a superar por el jugador, de cuyo diseño depende en gran parte la efectividad de una experiencia de juego: una curva muy abrupta provocará frustración en el jugador al serle imposible superar los retos propuestos; y una curva muy llana redundará en una experiencia aburrida al ser el juego poco desafiante y extremadamente simple de superar [7].

El juego consiste en la confrontación de las habilidades del jugador con el entorno virtual que le propone la narrativa, generalmente en forma de puzzles, trampas o enemigos. En los juegos más simples, el comportamiento de estos elementos es muy básico: los puzzles o trampas se encuentran en condiciones fijas, y los enemigos siguen movimientos preprogramados. Los motores de inteligencia artificial más elaborados, como los que son necesarios en un juego de peleas uno contra uno, o en un videojuego de fútbol, exigen algoritmos más complejos de toma de decisiones para lograr resultados más convincentes [8].

La teoría de juegos se define como el estudio de los modelos matemáticos del conflicto y la cooperación entre tomadores de decisiones inteligentes y racionales, la matemática de la decisión que permite un entendimiento cuantitativo de situaciones en las cuales dos o más individuos toman decisiones que afectarán al bienestar de cada uno de ellos [9]. Ésta nació en la década de 1940 como una disciplina aplicada a la economía y a la estrategia militar, y posteriormente su campo de aplicación se extendió a la biología, sociología, politología, psicología, filosofía y ciencias de la computación.

La teoría de juegos es de relevancia en las mecánicas de los videojuegos, en los cuales se desarrolla una confrontación entre los factores del videojuego y el/los jugador/es humano/s. La incorporación de algoritmos de teoría de juegos en los elementos de un serious game puede producir una experiencia más rica de usuario, ajustándose en forma dinámica de acuerdo a las condiciones del entorno y a las características del jugador. De hecho el balance entre las estrategias de un videojugador y el nivel de dificultad del juego, dado por las estrategias de los enemigos y los desafíos del mismo, son

críticos para una buena experiencia del jugador [10]. Los videojuegos comerciales, desde sus inicios con el mítico PacMan (NAMCO BANDAI Games, 1980), han incorporado algoritmos de toma de decisiones a sus elementos.

Por otra parte, la computación cuántica es una disciplina que aprovecha los efectos cuánticos en las partículas subatómicas para construir sistemas de computación capaces de resolver problemas de tiempo polinomial no determinista (NP) más eficientemente que sus contrapartes clásicos. Los problemas NP son aquellos de resolución fácil en una dirección y muy duros en la inversa, y que incrementan su complejidad en forma exponencial al aumentar linealmente el número de variables. La factorización de enteros, el problema del viajante de comercio o búsqueda de combinaciones para nuevas drogas, son algunos ejemplos de problemas NP [11].

La teoría de juegos cuánticos aplica el “esquema de cuantización” al campo de la teoría de juegos, añadiendo más flexibilidad y posibilidades a la teoría de juegos clásica.

El primer trabajo en combinar estos campos fue Quantum Strategies [12]. Más tarde ese mismo año, Eisert [13] definiría el método de cuantización para juegos que aún se aplica en la actualidad.

Meyer proporciona varios ejemplos de juegos de dos jugadores y suma cero, en los cuales se enfrentan un jugador restringido a algoritmos clásicos y el otro capaz de aplicar algoritmos cuánticos, y el jugador cuántico gana invariablemente. Estos ejemplos muestran las ventajas de las estrategias cuánticas sobre las clásicas para estos juegos, extensible a algoritmos de búsqueda y resolución de problemas. Eisert analiza el clásico “Dilema del Prisionero”, de la misma forma mostrando la superioridad de las estrategias cuánticas frente a las clásicas, dada la mayor riqueza de

movimientos que posee un jugador cuántico. La versión del Dilema del Prisionero con dos jugadores cuánticos aporta nuevas posibilidades al juego, aprovechando el efecto del entanglement.

Líneas de Investigación

- Diseño y Desarrollo de serious games.
- Teoría de juegos.
- Computación cuántica.

Resultados y Objetivos

El proyecto de tesis presentado en este artículo tiene como objetivo general el desarrollo de mecánicas novedosas en serious games, a partir de la aplicación de los principios de la teoría de juegos. Se considerará la incorporación de ramificaciones emergentes de esta teoría, tales como la teoría de juegos cuánticos, para contribuir en forma única y relevante a los campos del conocimiento involucrados.

Para ello se plantean los siguientes objetivos específicos:

- Explorar y analizar las bases conceptuales y metodológicas de la teoría de juegos, los algoritmos de computación cuántica y del desarrollo de serious games.
- Ponderar las posibilidades de aporte de los principios de las corrientes de teoría de juegos en las mecánicas de serious games.
- Desarrollar un conjunto de recursos de software que modelen un motor de inteligencia computacional basado en las tecnologías estudiadas.
- Incorporar un conjunto de mecánicas en serious games a partir de los recursos de software desarrollados.

- Evaluar el desempeño de las mecánicas creadas mediante metodologías de extracción de opinión de expertos.

Formación de Recursos Humanos

La formación de recursos humanos se considera de vital importancia.

En la actualidad, integrantes del Grupo de Investigación GTI se encuentran desarrollando y dirigiendo dos tesis de posgrado de la UNLP de la Maestría en Ingeniería de Software y dos de Doctorado, una en Educación de la Universidad Nacional de Rosario y otra del Doctorado en Ingeniería Orientación Modelado y Simulación Computacional de la UNMDP.

Referencias

- [1] Huizinga, J. (2000). *Homo Ludens*. Madrid, España: Ed. Alianza.
- [2] Abt, C. (1970). *Serious Games*. New York, Estados Unidos: Viking Press.
- [3] Zyda, M. (2005). From visual simulation to virtual reality to games. En *Computer* 38, pp. 25-32.
- [4] Sawyer, B., Rejeski, D. (2002). *Serious Games: Improving Public Policy through Game-based Learning and Simulation*.
- [5] Sawyer, B., Smith, P. (2008). *Serious games taxonomy*. En *Slides from the Serious Games Summit at the Game Developers Conference*.
- [6] Hamari, J., Koivisto, J. (2014). Measuring flow in gamification: Dispositional Flow Scale – 2. En *Computers in Human Behavior* (40).
- [7] Rogers, S. (2010). *Level Up: The Guide to Great Videogame Design*. John Wiley & Sons, Ltd. ISBN 978-0-470-68867-0
- [8] Buckland, M. (2004). *Programming Game AI by Example*. Jones & Bartlett Publishers, 1st edition .
- [9] Myerson, R. B. (1997). *Game Theory: Analysis of Conflict*. Boston, Estados Unidos: Harvard University Press.
- [10] Millington, I. (2006). *Artificial Intelligence for Games*. San Francisco, Estados Unidos: Morgan Kaufmann Publishers Inc..
- [11] Garey, M., Johnson, D. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. ISBN 0-7167-1045-5
- [12] Meyer, D. A. (1999). Quantum Strategies. En *Phys. Rev. Lett.* Volumen 82, pp. 1052-1055. American Physical Society.
- [13] Eisert, J., Wilkens, M., Lewenstein, M.(1999). Quantum Games and Quantum Strategies. En *Phys. Rev. Lett.*(83), pp. 3077. American Physical Society.

Analíticas de Aprendizaje para Serious Games

Stella Maris Massa¹, Franco Kühn²

Facultad de Ingeniería/ Universidad Nacional de Mar del Plata/Argentina
(7600) Av. Juan B. Justo 4302, +54-223-481660
smassa@fi.mdp.edu.ar¹, fdkuhn@gmail.com²

RESUMEN

La aplicación de juegos en la educación proporciona múltiples beneficios claramente visibles: su participación, orientada a la meta, anima a los estudiantes a progresar mientras se juega. Los videojuegos educativos denominados Serious Games son una iniciativa que se concentra en el uso de los principios de diseño de juegos para otros fines no meramente lúdicos.

Como educadores y desarrolladores de videojuegos, la validez de los Serious Games con respecto a sus objetivos educativos debe ser tanto mensurable como medida. La naturaleza interactiva de los videojuegos hace que la aplicación de Analíticas de Aprendizaje se constituya en una herramienta útil para capturar los datos de interacción de los estudiantes con el propósito de interpretar el proceso de aprendizaje. Sin embargo, existen escasas investigaciones y normas ampliamente adoptadas para comunicar información entre videojuegos y sus módulos de seguimiento.

El proyecto tiene como objetivo consolidar el proceso de análisis de la interacción de estudiantes en un Serious Game a través de la extracción de información relevante del aprendizaje y mediante la generación de reportes que visualicen los resultados.

Palabras clave: Serious Games, Interacción persona-computadora, Analíticas de Aprendizaje, Big-Data.

CONTEXTO

El proyecto que se presenta en este artículo corresponde a una Beca de investigación otorgada por la UNMDP.

Dicha Beca se desarrolla en el Grupo de Investigación en Tecnologías Interactivas (GTI) de la Facultad de Ingeniería de la Universidad Nacional de Mar del Plata.

Además las líneas de investigación se abordan en el marco del proyecto (2016-2017): “Tecnología e Innovación en Ambientes de Aprendizaje: Desarrollo y Gestión” y en el proyecto correspondiente a la convocatoria 2018-2019 ; “Modelos y herramientas para el proceso de desarrollo de Serious Games”. Ambos son financiados por la UNMDP.

1. INTRODUCCIÓN

Los Serious Games (SG) son aplicaciones interactivas creadas con una intencionalidad educativa, que proponen la explotación de la jugabilidad como experiencia del jugador. Presentan a los jugadores retos y misiones que implican tomas de decisiones, resolución de problemas, búsqueda de información selectiva, cálculos, exploración, experimentación, competencia, cooperación etc., logrando el efecto inmersivo en el juego, como una prolongación de la experiencia vital del usuario [1],[2].

La evaluación del aprendizaje en SG, lejos de ser sencilla, requiere métodos y modelos

adicionales que produzcan evaluaciones válidas y evidencias de aprendizaje basado en juegos, lo que requiere datos adicionales de los jugadores.

El informe anual Horizon [3], encargado de enumerar las tendencias educativas más influyentes del momento, señala a las Analíticas de Aprendizaje (Learning Analytics) como uno de los campos a tener en cuenta a mediano plazo. Las define como “el análisis de una amplia variedad de datos producidos por estudiantes, con el objetivo de evaluar su progreso académico, predecir sus futuros resultados e identificar riesgos potenciales”. En su edición de 2016, el Informe Horizon [4] destaca la medición de aprendizaje como uno de los campos donde se van a focalizar los mayores esfuerzos a corto plazo.

Las Analíticas de Aprendizaje constituyen la aproximación educativa a “Big-Data”, una ciencia originalmente utilizada en la empresa para analizar las actividades del consumidor, identificar tendencias de consumo y predecir comportamientos. En particular, la educación ha emprendido un proceso de búsqueda de datos similar, con el objeto de incrementar la retención de los estudiantes y proporcionarles una experiencia de aprendizaje personalizada y de alta calidad.

Tradicionalmente, la validación de los SG se ha hecho a través de experimentos formales mediante cuestionarios que los estudiantes completan antes y/o después de interactuar con el juego [5].

La riqueza actual de datos reunidos a través de la web-logs, motores de seguimiento, eye trackers, localización y detectores de movimiento, en combinación con las emergentes Analíticas de Aprendizaje es una excelente oportunidad para mejorar la supervisión y evaluación del aprendizaje basado en juegos.

Para aplicar esta tecnología emergente en los SG, los datos de la interacción pueden ser capturados mediante la adición de un “rastreador” al SG que envía datos de la interacción del jugador (también denominados rastreos) a un servidor. El análisis de las huellas puede producir información relevante sobre las interacciones de los estudiantes con el juego, haciendo que el conjunto de acciones, errores y acciones correctas del jugador sea significativo [6].

Coincidiendo con Alonso, et.al.[7], sin las Analíticas de Aprendizaje, los SGs en la educación son similares a las cajas negras: simplemente proporcionan un estado final que demuestra los resultados del juego, normalmente en forma de métricas relativas a la puntuación final del jugador; pero que no suministran información con respecto al proceso de aprendizaje.

Las Analíticas de Aprendizaje se han empezado a aplicar en múltiples medios y contextos para comprender y mejorar los procesos de aprendizaje: desde sistemas clásicos como los tutores inteligentes, hasta sistemas e-learning más modernos como los MOOCs. Sin embargo, las aproximaciones que se han construido sobre SG todavía no son maduras y no han hecho un uso completo de las ventajas que ofrecen dichas Analíticas.[8].

Las Analíticas de Aprendizaje en un SG comprenden desde los métodos de captura de datos y los elementos de diseño que los acompañan, hasta el análisis y visualizaciones más apropiados para cada tarea.

Para llevar a cabo esta integración se deben establecer algunos acuerdos básicos en las características del sistema de seguimiento de datos.

Tal como señala [9], la manera particular en que obtengan los datos determinará profundamente las posibilidades tanto en el

análisis como en la visualización, ya que implicará la cantidad de información que se recibirá y lo rica que es esa información. El análisis se desarrollará a partir de los datos obtenidos y sus particularidades. Las visualizaciones de resultados también deben ajustarse a las necesidades de los diferentes involucrados: estudiantes y profesores, como parte del uso educativo del juego, así como para desarrolladores o gestores, que estarían más interesados en el correcto funcionamiento del juego, pero también desean conocer si el juego está cumpliendo con los objetivos de aprendizaje previstos.

Es importante señalar que consideramos que las Analíticas de Aprendizaje en SG no sólo deben centrarse en el análisis o la visualización de datos, sino también deberían proveer a los profesores de herramientas que les permitan seguir el aprendizaje de sus estudiantes.

2. LINEAS DE INVESTIGACIÓN

Las principales líneas de investigación que se abordan en este proyecto son:

- Analíticas de Aprendizaje
- Evaluación del aprendizaje con tecnologías
- Serious Games y Gamificación

3. RESULTADOS ESPERADOS Y OBJETIVOS

El proyecto que se presenta en este artículo tiene como objetivo general: “Consolidar el proceso de análisis de la interacción de estudiantes en Serious Game para obtener información relevante del aprendizaje”.

En ese marco se establecieron los siguientes objetivos específicos:

- Seleccionar las tecnologías adecuadas para la visualización de las interacciones de los estudiantes en el Serious Game.
- Generar una estructura de visualización de las interacciones de los estudiantes en el Serious Game.
- Producir reportes que visualicen la información relevante del proceso de aprendizaje de los estudiantes sobre la base de sus datos de interacción en un Serious Game.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación GTI cuenta con varios integrantes que están realizando estudios de posgrado Maestrías y Doctorados. En particular, este artículo describe el plan de trabajo de una beca de investigación otorgada por la UNMDP.

Se dirigen además trabajos finales y prácticas profesionales supervisadas (PPS) de estudiantes de la Carrera Ingeniería en Informática.

5. REFERENCIAS BIBLIOGRÁFICAS

- [1] del Moral Pérez, M.E. (2013). Advergaming & edutainment: fórmulas creativas para aprender jugando. Ponencia inaugural del Congreso Internacional de Videojuegos y Educación (CIVE, 2013). Cáceres, España.
- [2] Westera, W., Nadolski, R., Hummel, H., Wopereis, I. (2008). Serious Games for Higher Education: a Framework for Reducing Design Complexity. *Journal of Computer-Assisted Learning* (Wiley), 24 (5), 420-432.
- [3] Johnson, L., Adams Becker, S. & Cummins, M. (2012). NMC Horizon Report: 2012 K-12

- Edition. Austin, Texas: The New Media Consortium.
- [4] Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., & Freeman, A., & Hall, C. (2015). Horizon Report: 2016 Higher Education Edition.
- [5] Calderón, A., & Ruiz, M. (2015). A systematic literature review on serious games evaluation: An application to software project management. *Computers & Education*, 87, 396–422.
- [6] Loh, C. & Sheng, D. & Ifenthaler, D.(2015), *Serious Games Analytics*. Springer International Publishing, Cham, 2015.
- [7] Alonso-Fernández, C., Calvo Morata, C. , Freire, M., Martínez-Ortiz, I. & Fernandez-Manjon, B. (2017). Systematizing game learning analytics for serious games, in: *IEEE Glob. Eng. Educ. Conf.*, 2017: pp. 1106–1113.
- [8] Serrano-Laguna, A. (2017). Mejorando la evaluación de juegos serios mediante el uso de analíticas de aprendizaje. Tesis Doctoral. Facultad de Informática. Universidad Complutense de Madrid.
- [9] Alonso- Fernández, C. (2016). Gaming learning analytics for serious games. Trabajo fin de Grado. Facultad de Informática. Universidad Complutense de Madrid.

Vinculación con docentes de nivel secundario y con ingresantes de Informática, a través de talleres y juegos educativos desarrollados, uno de los cuales, es un juego móvil.

Lic. Angela Belcastro ¹, APU. Macarena Quiroga ², APU. Juan Giménez ³,
 APU. Santiago Santana ⁴, APU. Pablo Dibez ⁵, Mg. Rodolfo Bertone ⁶
^{1, 2, 3, 4, 5} Ciencia y Técnica. Facultad de Ingeniería (FI). UNPSJB. Comodoro Rivadavia
⁶ III LIDI – Facultad de Informática – UNLP. (1900) La Plata, Buenos Aires, Argentina
¹(angelab)@ing.unp.edu.ar, ²msinf.quiroga@gmail.com, ³jm.gimenez.inf@gmail.com,
⁴santana.santiago@gmail.com, ⁵pdibez@gmail.com, ⁶pbertone@lidi.info.unlp.edu.ar

RESUMEN

La línea de investigación pertenece al área de **Tecnología Informática aplicada en Educación** (TIAE). Se desarrollaron tres **proyectos de investigación** (PI) continuados, y un cuarto proyecto se encuentra en etapa de evaluación inicial. Se ha generado, a través de ellos, vinculación con el medio, formación de recursos humanos de grado y postgrado, publicaciones con referato nacionales e internacionales, participaciones en congresos, propuestas didácticas para fomentar el **aprendizaje significativo** (AS), **juegos educativos interactivos** (JEI), organización, y dictado de talleres. Algunos de los integrantes de los PI son alumnos de “Licenciatura en Informática”, de la UNPSJB, sede Comodoro Rivadavia. El uso de las TICs en educación, el desarrollo de juegos educativos, son herramientas importantes, para generar motivación, incentivar a futuros profesionales de Informática, y fortalecer los estudios en los distintos niveles de enseñanza. Uno de los aspectos considerados en el nuevo proyecto, es el diseño y desarrollo de **sistemas interactivos de apoyo al aprendizaje** (SIAA) con **tecnología móvil** (TM), con el objeto de incentivar a los alumnos a investigar, y a participar en PI, de mejorar sus competencias, propiciando AS, fomentando el desarrollo de la sociedad y aumentando el conocimiento, formando personas competentes. El trabajo presenta algunos resultados de los talleres resientes.

Palabras Claves: TICs. Aprendizaje significativo. Tecnología móvil. Metacognición.

CONTEXTO

El proyecto de investigación acreditado: “Casos de estudio de sistemas, TICs y aprendizaje”, UNPSJB-UNLP, evaluado y financiado por Ciencia y Técnica de la UNPSJB, pertenece al Área TIAE. El proyecto continuado, titulado: “Tecnología Informática Aplicada en Educación, y Aprendizaje Significativo”, se encuentra en etapa de evaluación inicial, incluye el desarrollo de juegos educativos móviles, para motivar a los alumnos y propiciar AS.

1. INTRODUCCIÓN

Para **Piaget**, todo aprendizaje es el resultado de la interacción entre la persona y el objeto de conocimiento. Además, el **aprendizaje debe ser significativo**, ya que sólo así es capaz de modificar los esquemas de la persona. . Y para conseguir un **aprendizaje significativo**, es preciso favorecer la conexión entre las experiencias y conocimientos previos y los nuevos conocimientos. [1]
 Las **TICs** son esenciales en la educación, nos invitan a abordar el nuevo modelo de enseñanza y aprendizaje, propician la educación centrada en quien aprende. En los estándares de competencias en TICs propuestos por UNESCO, se afirma que: “Para vivir, aprender y trabajar con éxito en una sociedad cada vez más compleja, rica en información y basada en el conocimiento, los estudiantes y los docentes deben utilizar la tecnología digital con eficacia. En un contexto educativo sólido, las TICs pueden ayudar a los estudiantes a adquirir las capacidades necesarias”. La importancia del desarrollo de las **habilidades de “aprender a aprender”** se deriva no solo de concepciones teóricas propias del ámbito educativo, sino de las demandas sociales que parecen requerir la

formación de individuos capaces de un mayor manejo autónomo de estas **herramientas cognitivas**. [2]

“**Aplicaciones móviles**” es uno de los temas de mayor desarrollo en la industria del software y servicios informáticos, que ha generado cambios en los espacios curriculares universitarios en los últimos años.

Los **Juegos Educativos** ofrecen oportunidades para la resolución de problemas. Los **Juegos Educativos Móviles**, son un caso particular de Aplicación Educativa Móvil y éstos a la vez un caso particular de Aplicaciones Móviles. [3]

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El objetivo de esta línea de investigación, es el de contribuir a la formación de personas competentes, fomentar el desarrollo de la sociedad y aumentar el conocimiento, analizando y confeccionando propuestas educativas que propicien AS, diseñando y desarrollando SIAA, algunos de ellos, con tecnología móvil, con el objeto de incentivar a los alumnos a investigar, y a participar en PI, y mejorar las competencias de nuevos profesionales de Informática.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se han presentado otros resultados del proyecto, en congresos de Informática. Dentro de los juegos educativos interactivos, que se han desarrollado en el proyecto, encontramos juegos que fortalecen el pensamiento estratégico del usuario, con un nivel alto y medio, de dificultad. A principios de 2018 se han desarrollado dos encuentros, uno con alumnos ingresantes de Informática, y otro dirigido a docentes de nivel secundario. Estos talleres se desarrollaron en el marco de la Univerano 2018 de la UNPSJB, sede Comodoro Rivadavia.

El taller dirigido a ingresantes, nos permitió incentivarlos a participar en PI al avanzar en la carrera, participar en congresos, y capacitarse regularmente, también al obtener su graduación en la carrera, vinculándose con

universidades y/o institutos de investigación. Mostraron un alto nivel de motivación, al participar e interactuar con alumnos, disertantes, y con los juegos educativos.

En el taller dirigido a docentes, se realizó una exposición de los objetivos del encuentro, una presentación de cada participante y de los disertantes, y luego se organizaron las actividades, a través del siguiente esquema:



Figura 1. Esquema de trabajo del taller

Como vemos en la figura 1, las etapas del encuentro fueron seis, partiendo de una introducción inicial, en la que se resaltaron aspectos claves de las teorías de aprendizaje, los rasgos del buen aprendizaje, y la motivación por aprender, como factor relevante, para propiciar AS. Luego, se consideró la noción de software educativo, y se destacaron características básicas de los juegos educativos, y de los juegos educativos móviles. En la segunda etapa, se presentaron las características de los juegos que fortalecen el pensamiento estratégico, y cada estudiante de grado, que participa en proyecto, explicó el juego que desarrolló. En la siguiente etapa, los docentes participantes, interactuaron directamente con los juegos educativos. Luego completaron el formulario preparado para el taller. Dicho formulario, se tituló: “Docentes interesados en Juegos Educativos”, incluye inicialmente el objetivo, que es el de contar con información y mails de contacto de docentes interesados en:

- Aplicar en sus clases, alguno de los juegos educativos observados en el taller, resultantes de PI (con preguntas y respuestas preparadas por el docente que lo utilizará).
- Participar a largo plazo, en el diseño de un juego específico, de utilidad para sus clases, o de un software educativo.

- Comentar su experiencia en el taller, y en el uso de tecnología informática en el ámbito educativo.

La encuesta considera las siguientes dimensiones:

- “Docente”, en la que se solicita información de identificación del docente, de la institución de enseñanza, y materias principales que dicta.
- “Experiencia en el uso de las TICs”, en la que se invita a los participantes, a brindar comentarios acerca de los recursos educativos que ha utilizado en el aula, y además, de los recursos tecnológicos disponibles en el aula.
- “Interés en emplear juegos educativos observados en el taller”, en esta dimensión se indaga sobre el interés de uso de cada juego observado, con la posibilidad de responder:
 - “Si, en 2018”,
 - “No”,
 - “Puede ser más adelante”.
- “Nivel de satisfacción del taller desarrollado”, y un comentario adicional, si lo desea.

Los resultados indican que los docentes que participaron, eran docentes de nivel secundario, y una de las docentes, además, dictaba clases en la primaria, en sexto grado.

Todas las docentes participantes emplearon videos en sus clases. Una docente también utiliza comúnmente, libros digitales de su especialidad, y otra docente, ha construido un blog, y utiliza habitualmente los celulares en el aula con sus alumnos, para búsquedas de información.

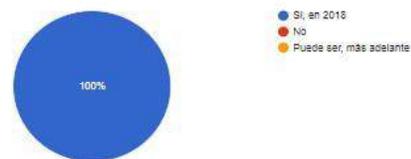
Dos de las docentes destacan que la mayoría de los alumnos tienen celulares en las clases. Una docente señala que la mitad de los alumnos, en la clase, tienen celulares. Las instituciones cuentan con cañón de proyección. Una docente destaca que en la institución en la que desarrolla actividades, disponen de sala de computación con una computadora para cada alumno. Otra docente indica que la sala de computación de la escuela en la que ejerce su actividad docente, tiene computadoras para la clase, trabajando dos alumnos en cada computadora. Una

docente indica que la conexión a Internet es muy lenta, en el establecimiento educativo.

Dos docentes están interesadas en emplear en sus clases, en 2018, los juegos 1 y 2. Una docente indica que puede tener interés, más adelante en usar los juegos 1 y 2. Estos juegos son variantes del Juego de Símbolos, uno de ellos permite el entrenamiento del alumno, en sistemas binarios.

C.3- ¿Tenés interés en emplear en tu actividad docente el juego 3
Juego del Solitario con Monedas ?

3 respuestas



C.4- ¿Tenés interés en emplear en tu actividad docente el juego 4
Juego Móvil El ahorcado ?

3 respuestas



Figura 2. Interés de uso de juegos, en 2018

Como muestra la figura 2, todos los docentes participantes destacaron su interés en emplear, en 2018, los juegos 3 (Juego del Solitario con Monedas) y 4 (Juego Móvil El Ahorcado). Dichos juegos se consideraron apropiados tanto para nivel secundario, como para sexto año, de nivel primario.

Las tres docentes participantes mostraron interés en integrar un proyecto en el que se diseñe un juego interactivo específico, que sea de utilidad para sus clases. Dos de los docentes participantes, consideraron, en una escala del cero al 4, 4, en el nivel de satisfacción con la experiencia, y una docente consideró 3, su nivel de satisfacción. Los comentarios adicionales de los docentes, son:

- Sería bueno que los alumnos se acerquen de manera diferente a los juegos interactivos, aportando un plus de conocimiento que puede ser aplicado a lo cotidiano.
- Si se concreta la idea, aplicarla en una clase en mi escuela para comprobar y ponerlo en práctica.

Los integrantes de proyecto realizarán este mes, otros talleres ya programados, con estudiantes y docentes, con análisis de uso de los dispositivos y transferencia, en distintos niveles de enseñanza.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto de investigación acreditado: “Casos de estudio de sistemas, TICs y aprendizaje”, UNPSJB-UNLP, evaluado y financiado por Ciencia y Técnica de la UNPSJB. El proyecto continuado, titulado: “Tecnología Informática Aplicada en Educación, y Aprendizaje Significativo”, se encuentra en Ciencia y Técnica, en etapa de evaluación inicial. El director del PI es docente-investigador, de la UNLP, categoría II en investigación, la codirectora es docente-investigadora de la UNPSJB, categoría III en investigación, el equipo está integrado por docentes de la Licenciatura en Informática, y profesionales Analistas Programadores Universitarios, recibidos en la UNPSJB, que desarrollan actividades de Informática en empresas de Comodoro Rivadavia. El equipo completo tiene siete integrantes. Todos los integrantes se capacitan regularmente, y avanzan, en carreras de grado, o en carreras de postgrado.

5. BIBLIOGRAFIA

- [1] Begonia Gros Salvat. Aprendizajes, conexiones y artefactos. La producción colaborativa del conocimiento. Gedisa editorial. 2008.
- [2] Camilloni, A. R. Celman, S. Litwin, E. y Palou de Maté M. “La evaluación de los aprendizajes en el debate didáctico contemporáneo”. PAIDÓS Educador. 2010.
- [3] Lliteras Alejandra Beatriz. Gordillo Silvia. Una Guía para la Conceptualización de Juegos Educativos Móviles. UNLP. 2013.
- [4] Calero Pérez, M. Constructivismo pedagógico. Teorías y aplicaciones básicas. Alfaomega. 2008.
- [5] Mg. GONZÁLEZ DE DOÑA y otros. Desarrollo de competencias para la gestión de información y construcción de conocimientos: TICs y Nuevos Ambientes Educativos. Departamento de Informática/ Programa Permanente de EAD/ Facultad de Ciencias Exactas, Físicas y Naturales/ Universidad Nacional de San Juan. 2009.
- [6] Horacio Ferreyra. Griselda Gallo. Ariel Zecchini. Educar en la acción para aprender a emprender. Organización y gestión de proyectos socio-productivos y cooperativos. Noveduc. 2013.
- [7] Asinsten, G., Espiro, M. S. y Asinsten. J. Construyendo la clase virtual. Métodos, estrategias y recursos tecnológicos para buenas prácticas docentes. Novedades Educativas. Didáctica. 2012.
- [8] Ritchhart, R., Church, M. y Morrison, K. Hacer visible el pensamiento. Cómo promover el compromiso, la comprensión y la autonomía de los estudiantes. PAIDÓS. 2014.
- [9] Hernández Forte, V. Mapas conceptuales. La gestión del conocimiento en la didáctica. Alfaomega. 2005.
- [10] Anijovich, R. Camilloni, A. R. Cappelletti, G. Hoffmann, J. Katzkowicz, R y López, L. M. La evaluación significativa. PAIDÓS. 2010.
- [11] Gustavo Alfredo Bacino. Aula extendida en la educación superior en ingeniería. Una propuesta de aplicación en el Área Tecnológica Básica de Electrotecnia. 2014. Acceso agosto 2017. <http://sedici.unlp.edu.ar/handle/10915/45476>
- [12] Santiago Castillo Arredondo. Luis Polanco González. Enseñar a estudiar... aprender a aprender. Didáctica del Estudio. Pearson. Prentice Hall. 2005.
- [13] Belcastro A. y otros. Técnicas de pensamiento y Teoría Triárquica como elemento para el desarrollo de software educativo. 2008.
- [14] Belcastro A. y otros. EduIAS, como estrategia de formación que apunta a lograr un aprendizaje significativo y colaborativo asistido por computador. 2005.
- [15] Litwin E. El oficio de enseñar. Condiciones y contextos. PAIDÓS. 2008.

- [16] Andrés Senlle. Nilda Gutierrez. Calidad de los servicios educativos. Diaz de Santos. 2005.
- [17] Cifuentes, R. M. Formulación de proyectos pedagógicos. Para mejorar la enseñanza universitaria. Fundamentación – Redacción – Evaluación. Colección Universidad. Noveduc. 2014.
- [18] Maggio, M. Enriquecer la enseñanza. Los ambientes con alta disposición tecnológica como oportunidad. PAIDÓS. 2012.
- [19] Mg. Roberto Bertone. Mg. José Luis Filippi. Lic. Guillermo Lafuente. Mg. Carlos Ballesteros. Lic. Gustavo Lafuente. I.S. Daniel Perez. I.S. Sofía Aguirre. Tecnología Móvil aplicada en la educación. 2015. Acceso en agosto de 2017. <http://hdl.handle.net/10915/46269>
- [20] Susana I. Herrera. Marta C. Fennema. Tecnologías Móviles Aplicadas a la Educación Superior. Universidad Nacional de Santiago del Estero. Universidad Tecnológica Nacional, Facultad Regional Tucumán. CACIC 2011. Acceso en agosto de 2017. <http://hdl.handle.net/10915/18718>
- [21] Leda Digión. Marisa Digión. Mabel Sosa. Aprendizaje innovador con uso de tecnología móvil. 2014. Acceso en agosto de 2017. <http://hdl.handle.net/10915/43694>
- [22] Alfonso García García. Roger Rey Barbáchano. Las apps en el aula del siglo XXI. Centro de Comunicación y Pedagogía. 2014. Acceso agosto 2017. Disponible en: <http://www.centrocp.com/las-apps-en-el-aula-del-siglo-xxi/>
- [23] Joo Nagata, J. & Martínez Abad, F. (2015). Patrimonio territorial virtual en educación: recursos en mlearning sobre la ciudad de Salamanca. En AIDIPE (Ed.), Investigar con y para la sociedad (Vol. 3, pp. 1681-1692). Cádiz, España: Bubok. Acceso agosto 2017. Recuperado de <http://aidipe2015.aidipe.org>
- [24] Carly Shuler. Niall Winters. Mark West. UNESCO. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. El futuro del aprendizaje móvil. Implicaciones para la planificación y formulación de políticas. Edición y diseño gráfico: Rebecca Kraut. ISSN 2305-8617. 2013.
- [25] UNESCO. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. Directrices para las políticas de aprendizaje móvil. ISBN 978-92-3-001145-1. 2013.
- [26] Morales Caluña. Edgar Rolando. Diseño de objetos móviles de aprendizaje para mejorar el rendimiento académico de los estudiantes en la cátedra de Programación IV en el Instituto Tecnológico Pelileo. 2014. Acceso agosto de 2017. <http://dspace.esepoch.edu.ec/handle/123456789/4136>
- [27] Matías Loto. Elena Durán. Diseño de una aplicación móvil personalizada de apoyo al aprendizaje de Redes de Computadoras. Instituto de Investigaciones en Informática y Sistemas de Información (IISI), Facultad de Ciencias Exactas y Tecnologías (FCEyT) Universidad Nacional de Santiago del Estero (UNSE). 2015. Acceso en agosto de 2017. <http://hdl.handle.net/10915/49032>
- [28] Javier Cuello. José Vittone. Diseñando apps para móviles. 2013. Acceso agosto 2017. <https://books.google.com.ar/books?id=ATiqsjH1rvwC&printsec=frontcover&hl=es#v=onepage&q&f=false>
- [29] Durán Sanjuán, Álvaro Javier, Peinado Rodríguez, Jorge Luis, Rosado, Albeiro Alonso, Comparación de dos tecnologías de desarrollo de aplicaciones móviles desde la perspectiva del rendimiento como atributo de calidad. Scientia Et Technica [en línea] 2015, 20 (Marzo-Sin mes). Acceso agosto 2017. Disponible en: <<http://www.redalyc.org/articulo.oa?id=84938609010>> ISSN 0122-1701
- [30] Ministerio de Producción. Ministerio de Educación y Deportes. Programa 111 MIL. Analistas del Conocimiento. Dimensión Programador. Apunte del Módulo. Desarrollo de Software. Version 1.1. Liberada 08-02-2017.

Realidad virtual en terapias para estimular la cognición de niños con encefalopatía crónica no evolutiva

Cintia Ferrarini ¹, Mónica González ², Luís Olguín ¹, Emilio Ormeño ¹

¹ Instituto de Informática; Facultad de Ciencias Exactas, Físicas y Naturales; UNSJ

² Departamento de Informática; Facultad de Ciencias Exactas, Físicas y Naturales; UNSJ

¹ {eormeno, ferrarini, lolguin, szapata}@iinfo.unsj.edu.ar, ² gonzalez.monica@gmail.com

1. RESUMEN

Los niños con encefalopatía crónica no evolutiva (ECNE) poseen diversas dificultades motoras y cognitivas que los limitan para interactuar en forma autónoma con su entorno; y es justamente esa autonomía la que se intenta propiciar en las intervenciones terapéuticas debido a que ella impacta en forma directa en su calidad de vida. En el proyecto “*Videojuegos en realidad virtual como intervención terapéutica alternativa para estimular la cognición en niños con encefalopatía crónica no evolutiva*” (Vivitan), se pretende desarrollar un videojuego de realidad virtual (VR) para niños con ECNE que ofrezca un entorno inmersivo para la realización de actividades lúdicas motivantes que estimulen sus funciones cognitivas: atención, concentración y memoria; mientras ejecuta movimientos controlados con la cabeza. La tecnología desarrollada tendrá como destinatarios a los niños y adolescentes del Centro de Rehabilitación Integral, organismo estatal de la salud que brinda intervenciones terapéuticas a personas con discapacidad entre otras actividades, y que se constituye como la institución demandante.

2. CONTEXTO

Vivitan, se enmarca en “*Proyectos de Desarrollo Tecnológico y Social*” (PDTs), correspondiente a la convocatoria 2018-2019

y en la línea de investigación “*Videojuegos para a la salud*”. En esta línea el equipo de investigadores viene trabajando desde el año 2013, en el proyecto de investigación denominado “*Impulso a la Producción de Videojuegos Aplicados a Salud*” (IPVÍAs), financiado por la Secretaría de Políticas Universitarias (SPU) según Resolución SPU-2919-2013. También a través de los proyectos “*Videojuegos para motivar conductas Saludables*” y “*Mover-T*” aprobados por CICITCA-UNSJ en el marco de las convocatorias 2014–2015 y 2016-2017 respectivamente.

Vivitan se llevará a cabo en el Instituto de Informática de la Facultad de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de San Juan, en una tarea conjunta con la institución demandante para la generación de una tecnología apropiada.

3. INTRODUCCIÓN

La encefalopatía crónica no evolutiva (ECNE), es una patología que produce una discapacidad que se produce con causas prenatales, natales o postnatales hasta los tres años. Siempre predomina una alteración motora y puede estar asociada a una alteración visual, auditiva, intelectual del lenguaje, etc.

La realidad virtual (VR) es un escenario generado por computadora que simula una experiencia realista. El entorno inmersivo puede ser similar al mundo real para crear una

experiencia realista basada en la realidad o la ciencia ficción. Los sistemas de realidad aumentada también se pueden considerar una forma de realidad virtual que coloca información virtual sobre una cámara en tiempo real, en un casco de VR, o a través de un teléfono inteligente o tableta.

La VR es cada vez más aceptada por la sociedad debido a la creciente disponibilidad de sus dispositivos y aplicaciones. Muchos videojuegos en VR proponen entornos inmersivos de aprendizaje basados en el descubrimiento y en la creatividad. Los mismos, debido a su naturaleza lúdica, suponen un mecanismo de estímulo para el desarrollo cognitivo porque facilita el proceso de aprendizaje, incrementan la atención y la concentración a la hora de resolver un problema concreto.

En el ámbito de la salud, muchos trabajos han producido resultados aceptables utilizando VR en el tratamiento de patologías discapacitantes entre las que se encuentra la encefalopatía crónica no evolutiva, antes conocida como “parálisis cerebral”.

El gran desafío tecnológico de la VR aplicada a videojuegos para niños con ECNE se centra en lograr que realicen actividades lúdicas que estimulen su cognición (atención, concentración y memoria) ejecutando movimientos controlados de la cabeza. El objetivo final es que el juego le ayude al niño a desarrollarse y divertirse como otros niños. Tal como dice Vigotsky: *“Las funciones mentales superiores se adquieren y se desarrollan a través de la interacción social y el juego es un gran agente socializador”*.

El Proyecto Vivitan es un proyecto de investigación aplicada que se sustenta en las siguientes hipótesis:

1. Los videojuegos en Realidad Virtual pueden estimular la atención, la concentración y la coordinación visomotriz en niños con ECNE.
2. El proceso de desarrollo de videojuegos en este contexto requiere adaptaciones.

En el marco de esta investigación se ha optado por un diseño descriptivo experimental y por una metodología cuantitativa. Para el desarrollo del software se utilizará una metodología ágil basada en Scrum. Las instancias de evaluación de usabilidad se realizarán con usuarios reales de la institución demandante.

Se espera que dicho desarrollo tecnológico contribuya a favorecer las capacidades cognitivas de atención, concentración y memoria en niños con ECNE.

4. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

El proyecto Vivitan se encuadra dentro de los procesos de Ingeniería de Software aplicados a videojuegos para la salud y la educación; y dentro del Diseño de Videojuegos para personas con discapacidad.

5. RESULTADOS ESPERADOS/OBTENIDOS

El conocimiento en la temática de videojuegos para la salud, obtenido en el marco de los proyectos predecesores a Vivitan, permitió el desarrollo de prototipos de videojuegos destinados a la rehabilitación motriz de niños con ECNE. Los que fueron puestos a disposición de algunos usuarios para su uso. Esto permitió conocer mejor las características de los usuarios destinatarios de la tecnología y proponer mejoras y ajustes en los prototipos. Esta experiencia adquirida es

la que servirá de base para el proyecto Vivitan..

Para la ejecución del mismo, se propone identificar las funciones ejecutivas y capacidades del grupo destinatario mediante la administración de entrevistas focalizadas al equipo de terapeutas.

Los sujetos bajo estudio con los que se trabajará en Vivitan son niños y adolescentes que asisten al Centro de Rehabilitación Integral, cuyas edades cognitivas se encuentran comprendidas entre 10 y 16 años.

Durante la ejecución del proyecto se adaptará una metodología de desarrollo de software ágil que contemple las características del videojuego. Específicamente se utilizará la metodología SUM para desarrollo de videojuegos.

Vivitan responde a una de las necesidades priorizadas del Centro de Rehabilitación Integral, debido a las escasas terapias alternativas con que cuenta para apoyar las intervenciones terapéuticas en niños con ECNE.

A fin de explorar nuevos enfoques en la mejora de sus aptitudes cognitivas tales como la concentración, la memoria, y la resolución de problemas, se pretende propiciar espacios para que los terapeutas manejen nuevas herramientas como videojuegos en realidad virtual.

6. FORMACIÓN DE RRHH

1. Ormeño, Emilio Gustavo. Director del proyecto. Actualmente doctorando del Doctorado en Ciencias de la Informática, y cuya tesis está orientada al área de Videojuegos para la salud.
2. González, Mónica Gilda. Co-directora. Actualmente desarrollando su tesis Doctoral en Educación, en instancia de Informe Final, Doctorado en Educación de la Universidad Católica de Cuyo, Título Procesos de Apropiación de TIC en Espacios Socio-educativos no Formales. Propuesta de Formación Integral para Adultos Mayores.
3. Ferrarini Oliver, Cintia. Integrante del proyecto. Actualmente doctorando del Doctorado en Ingeniería de la Universidad Nacional de Cuyo. Título Método de Planificación Estocástico basado en Markov para la Composición de Servicios Web en un Entorno de Recomendación.
4. Olguín, Luis Alberto. Integrante del proyecto. Actualmente maestrando de la Maestría en Informática de la Universidad Nacional de San Juan. Título Red de Co-Préstamo en Bibliotecas”.
5. Vera Cristina. Actualmente maestrando de la Maestría en Informática de la Universidad Nacional de San Juan, con el título Estudio comparativo del desempeño de diferentes gestores de bases de datos NoSQL en la nube.
6. Avendaño Mauro. Trabajo Final de Licenciatura en proceso. Título Mover-TX: Gestor de recursos distribuidos de videojuegos en línea basados en Unity.
7. Carrió Fabricio. Trabajo Final de Licenciatura en proceso. Título Dungeon World: juego de rol en 3D desarrollado en plataforma Roblox.
8. Kokot Rodrigo: Trabajo Final de Licenciatura en proceso. Título Obtención de parámetros corporales con Kinect

9. Morales Rubén: Actualmente maestrando de la Maestría en Informática de la Universidad Nacional de San Juan, con el título Una metodología para el desarrollo de classroom board games conservando la experiencia de usuario.

7. BIBLIOGRAFÍA

Diez-Alegre, M. I., & Muñoz-Hellín, E. (2013). Empleo de sistemas de realidad virtual sobre la extremidad superior en niños con parálisis cerebral. Revisión de la literatura. *Fisioterapia*, 35(3), 119–125. <http://doi.org/10.1016/j.ft.2012.10.002>

Monge Pereira, E., Molina Rueda, F., Alguacil Diego, I. M., Cano de la Cuerda, R., de Mauro, A., & Miangolarra Page, J. C. (2014). Empleo de sistemas de realidad virtual como método de propiocepción en parálisis cerebral: guía de práctica clínica. *Neurología*, 29(9), 550–559. <http://doi.org/10.1016/j.nrl.2011.12.004>

Pirila, Silja, van der Meere, J., Korhonen, P., RuusuNiemi, P., Kyntaja, M., Nieminen, P. y Korpela, R. (2004). A retrospective neurocognitive study in children with spastic diplegia. *Developmental neuropsychology*, 26(3), 679-690. doi:10.1207/s15326942dn2603_2.

Bottcher, L. (2010). Children with Spastic Cerebral Palsy, Their Cognitive Functioning, and Social Participation: A Review. *Child Neuropsychology*, 16(3), 209-228. doi:10.1080/09297040903559630.

Rai, Y., Chaturvedi, S., Paliwal, V. K., Goyal, P., Chourasia, A., Singh Rathore, R. K., Gupta, R. K. (2013). DTI correlates of cognition in term children with spastic diplegic cerebral palsy. *European journal of*

paediatric neurology: EJPN: official journal of the European Paediatric Neurology Society, 17(3), 294-301. doi:10.1016/j.ejpn.2012.11.005.

Yagüe Sebastián, M. P., Yagüe Sebastián, M. M., Lekuona Amiano, A., & Sanz Rubio, M. C. (2016). Los videojuegos en el tratamiento fisioterápico de la parálisis cerebral. *Fisioterapia*, 38(6), 295–302. <http://doi.org/10.1016/J.FT.2015.11.005>

SUM para el desarrollo de videojuegos. (n.d.). Retrieved March 09, 2018, from <http://www.gemserk.com/sum/>

Discapacidad y salud. (OMS). Retrieved March 09, 2018, from <http://www.who.int/mediacentre/factsheets/fs352/es/>

Aplicación de las Reglas de Usabilidad en Sitios Web Universitarios.

Olariaga, Sandra ¹; Ligorria, Laura ²; Di Gionantonio, Alejandra ³;
Carrizo, Blanca ⁴; Colacioppo, Nicolás ⁵; Peralta, Lorena ⁶; Scandolo Carlos Iván ⁷;
Manera, Roxana ⁸; Savorgnano, Marcela ⁹

Laboratorio de Investigación de Software (Lis)
Departamento de Ingeniería en Sistemas de Información
Universidad Tecnológica Nacional, Facultad Regional Córdoba (UTN-FRC)
Maestro Marcelo López y Cruz Roja Argentina – Ciudad Universitaria – Córdoba, Argentina

⁽¹⁾solariaga@yahoo.com; ⁽²⁾liuniversidad@gmail.com; ⁽³⁾ing.alejandrag@gmail.com;
⁽⁴⁾bcarrizo@frc.utn.edu.ar; ⁽⁵⁾nicolas_colacioppo@hotmail.com
⁽⁶⁾peralta.lorena.d@gmail.com;
⁽⁷⁾ivansc@bs.frc.utn.edu.ar; ⁽⁸⁾r_manera@hotmail.com;
⁽⁹⁾marale511@gmail.com;

Eje Temático del Trabajo: Tecnologías informáticas aplicadas a la educación.

RESUMEN

En el contexto de la Ingeniería de Software entre las disciplinas que componen la misma mencionaremos Validación y Verificación de Software, la cual se encarga de efectuar el Testing del mismo. En el estudio de campo se tuvo en cuenta la Interfaz de usuario de sitios web universitarios, desde el momento en que la página esté funcional. En breve estaremos probando una herramienta elaborada que contempla los diferentes aspectos que involucran la evaluación de la usabilidad de los sitios web universitarios seleccionados. Este trabajo tendrá una incidencia directa en los usuarios de sitios web universitarios, ya que como resultado del proceso de investigación se generará un conjunto de buenas prácticas en la construcción de portales universitarios, lo cual facilitara el acceso a los usuarios a dichos sitios web universitarios.

Palabras clave: usabilidad, interface de usuario, sitios web universitarios, herramienta.

CONTEXTO

El presente proyecto está siendo llevado a cabo en el Laboratorio de Investigación de Software, Departamento de Ingeniería en Sistemas de la Universidad Tecnológica Nacional Facultad Regional Córdoba. (U.T.N.-F.R.C.).

El Código del proyecto es UTN4082 dentro del área de Ciencias de la Computación e Informática, radicado dentro de la Secretaria de Ciencia y Tecnología de la Universidad Tecnológica Nacional Facultad Regional

Córdoba. Fecha de inicio: 01/05/2016 Fecha de finalización: 31/12/2018.

1. INTRODUCCIÓN

Este trabajo es la continuación del iniciado el año pasado [Sandra Olariaga et al, 2017]. Las siguientes son las definiciones de usabilidad presentada:

"La usabilidad trata sobre el comportamiento humano; reconoce que el humano es emotivo, no está interesado en poner demasiado esfuerzo en algo, y generalmente prefiere las cosas que son fáciles de hacer contra las que son difíciles de hacer." [David Mc Quillen, 2003].

Usabilidad es un término adaptado de la palabra en inglés "usability", para indicar que algo se puede usar, la norma ISO 9241-11 dice que la usabilidad se refiere al alcance en el que un producto puede ser utilizado por usuarios específicos para alcanzar metas específicas con efectividad, eficiencia y satisfacción en un contexto específico de uso. [Guillermo M. Martínez de la Teja, 2003].

En los Sistemas de Información Web existen conjuntos de atributos que son interesantes medir, como son: la usabilidad, la accesibilidad, la navegación, etc. Estos atributos son evaluados una vez que el sistema es desarrollado. [Aguirre, Riesco, Figuereido, 2012]

En base a lo mencionado anteriormente, se hace evidente que la usabilidad juega un papel primordial en el proceso de desarrollo de sitios web de éxito. [Perurena Cancio et al, 2013]. En el presente trabajo, se pretende continuar con la identificación de los aspectos a ser tenidos en cuenta para el desarrollo de una herramienta con la cual se encuestaran usuarios cuyos perfiles fueron definidos en este trabajo, a fin de obtener información acerca de la usabilidad de interface de usuario de sitios web universitarios nacionales.

Para ello teniendo en cuenta además su arquitectura [Montes de Oca Sánchez de Bustamante, Antonio (2004)], es decir su estructura y organización ya que aunque para la mayoría de los usuarios "la interface es la aplicación" puesto que es la parte que ven y a través de la cual interactúan [Mario Moreno, et al, 2008], debemos entender que la usabilidad de la aplicación depende no sólo del diseño del interfaz, sino también de su arquitectura es decir del componente no visible del diseño. [Yusef Hassan, Francisco J. Martín Fernández y Ghzala Iazza 2004].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto se inscribe dentro de los lineamientos de investigación en innovaciones curriculares en Educación

Superior Universitaria dentro del área de Sistemas e interfaces adaptables en la intervención humano-computadora.

Para el desarrollo de este trabajo de investigación se aplicará el método empírico-analítico, que se basa en la experimentación y en la lógica empírica, junto a la observación de sitios web universitarios y análisis de usabilidad de los mismos

3. RESULTADOS OBTENIDOS / ESPERADOS

El eje sobre el que gira nuestro trabajo es la interface de usuario centrada en la evaluación de los sitios web universitarios nacionales. [Paula Maciel (2008)].

Dicha evaluación está basada en el análisis de la interacción de los diferentes tipos de usuarios con estos sitios web.

El principal objetivo de una interface de usuario es que esta pueda comunicar información a través de ella hacia algún tipo de dispositivo o sistema. Conseguida esta comunicación, el siguiente objetivo es que dicha comunicación se desarrolle de la forma más fácil y cómoda posible para las características del usuario que utiliza el servicio. Por ello este trabajo se enfoca en lograr identificar las técnicas que nos permitan realizar una evaluación eficaz de la

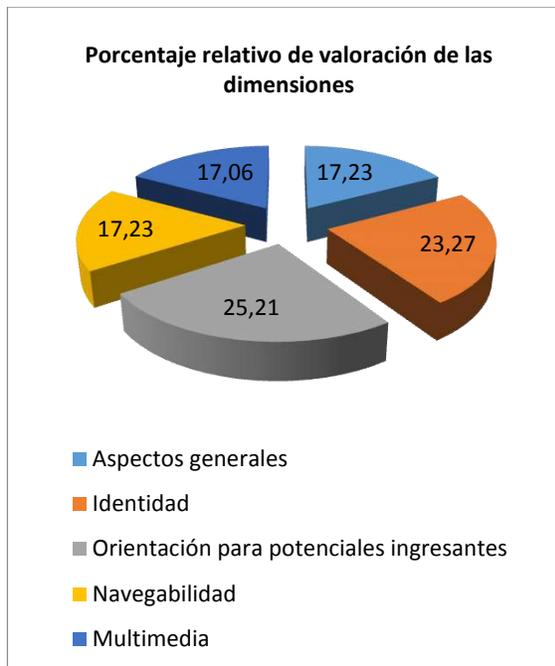
interface de usuario de los sitios web universitarios propuestos.

En base al análisis de los principales aspectos a ser tenidos en cuenta sobre el diseño de los sitios web, a través de un detallado análisis de la bibliografía existente hemos llegado a la siguiente tabla de dimensiones y subdimensiones:

Dimensión	Subdimensión
Aspectos generales	Claridad
	Sencillez
	Versatilidad
	Medios de transporte
Identidad	Claridad
	Medios de contacto
	Logotipo
Orientación para Potenciales ingresantes	Becas
	Trámites administrativos
	Volumen de información
	Tutorías
Navegabilidad	Mapa del sitio
	Sentido de avance
	Sencillez
	Búsqueda
	Manejo de errores
Multimedia	Tamaño de imágenes
	Tipo de imágenes

Tabla de Dimensiones y Subdimensiones

A continuación se muestran los resultados de una prueba piloto llevada a cabo, en este caso la evaluación de la página web del Instituto Tecnológico Buenos Aires.



Este gráfico muestra el porcentaje relativo de valoración del sitio web del Instituto Tecnológico de Buenos Aires en las cinco diferentes dimensiones evaluadas en las encuestas. Se puede apreciar que las dimensiones de orientación para potenciales ingresantes y de identidad, son las mejores valoradas. La implementación de estas pruebas piloto, consisten en encuestas a diferentes usuarios, a fin de poder ajustar la herramienta de recolección de datos, de manera de obtener respuestas para elaborar el conjunto de recomendaciones y buenas prácticas aplicables en el diseño y construcción de los sitios web universitarios

4. FORMACIÓN DE RECURSOS HUMANOS

El desarrollo del Proyecto que engloba este trabajo, tendrá un alto impacto en la formación de RRHH en el área de desarrollo web mediante la determinación de pautas para lograr la eficiencia y la optimización de la usabilidad del diseño de los sitios web universitarios. Los investigadores que recién inician forjarán sus primeras herramientas en I+D lo que les permitirá ampliar conocimientos en el área en cuestión además de incrementar sus antecedentes académicos. El proyecto contribuirá también a consolidar los conocimientos de dos integrantes del grupo, uno que está realizando la Diplomatura en Testing; y a fortalecer la formación de un becario, estudiante de Ingeniería en Sistemas de Información. A su vez induce a la interacción de trabajo en conjunto de investigadores de diferentes ámbitos y departamentos académicos de la Facultad Regional Córdoba tales como el Departamento de Ingeniería en Sistemas de Información y el Departamento de Ciencias Básicas que unirán sus esfuerzos logrando un intercambio de experiencias y conocimientos que enriquecerán su saber.

5. BIBLIOGRAFÍA

-Aguirre, Riesco, Figueredo, “Evaluación de la Usabilidad en el desarrollo de Sistemas Web dirigido por modelos” WICC 2012

-Carlos González, Iván Darío, Collazos Cesar. (2006) Propuesta Metodológica para la Evaluación de la Usabilidad en Sitios Web: Experiencia Colombiana Departamento de Telemática Universidad del Cauca FIET, Sector Tulcán, Popayán, Colombia

-David, McQuillen, “Taking Usability Offline”, Darwin Magazine, June 2003.

-González Carlos - Evaluación de calidad web: Métodos, técnicas y uso de métricas de usabilidad (2016)

http://www.usabilidadweb.com.ar/metodos_eval_calidad_web.php

-Guillermo Martínez de la Teja, “Usabilidad y Accesibilidad en WEB”. Sociedad de Ergonomistas de México. Universidad. Fecha de consulta: 01-08-2016. Disponible en <http://www.semec.org.mx/archivos/6-11.pdf>

-Luna Huertas Paula, Aguaded Gómez José, Martínez López Francisco., García Ordaz Mercedes, Martínez López Antonio, Pardo Rojas Antonio (2004) Hacia una metodología de análisis de sitios web universitarios EDUTEC Barcelona España.

-Mario Moreno, et al. “Evaluación de la Calidad en uso de sitios web asistida por Software: SW-AQUA” México 2008.

-Montes de Oca Sánchez de Bustamante, Antonio (2004) Arquitectura de información y usabilidad tomado de: http://www.wikilearning.com/usabilidad_nociones_basicas_para_los_profesionales_de_la_informacion-wkccp-8135-1.htm http://bvvs.sld.cu/revistas/aci/vol12_6_04/aci04604.htm

-Lilliam Perurena Cancio, Mercedes Moráguez Bergues. (2013). Usabilidad de los sitios Web, los métodos y las técnicas para la evaluación. Revista Cubana de Información en Ciencias de la Salud. ISSN 2307-2113. Vol. 24, Núm. 2. Disponible en: <http://www.acimed.sld.cu/index.php/acimed/article/view/405/306>

-Olariaga Sandra, et al. (2016) “Interface de sitios web universitarios: análisis de usabilidad” II JATIC. II Jornadas Argentinas de Tecnología, Innovación y Creatividad 2016. <http://www.jatic2016.ucaecemdp.edu.ar/>

-Paula Maciel (2008) Análisis de la eficacia de los sitios web universitarios argentino, Carrera de Maestría en Educación, Tesis de Maestría. Bella Vista Buenos Aires. <http://www.caminandoutopias.org.ar/tesis/tesis11.php>

Yusef Hassan, Martin Fernández Francisco y Lazza Ghzala (2004) Diseño Web Centrado en el Usuario: Usabilidad y Arquitectura de la Información Universidad de Granada Hipertext.net https://www.upf.edu/hipertextnet/numero-2/disenos_web.html

MODELO PREDICTIVO PARA LA DETERMINACIÓN DE RIESGO DE DESERCIÓN UNIVERSITARIA PARA INGRESANTES DE LA UTN FRLP

ISTVAN, Romina; LASAGNA, Valeria

*Universidad Tecnológica Nacional, Facultad Regional La Plata
Laboratorio de Ingeniería en Sistemas de Información, LINES
Av. 60 s/n° esquina 124, CP 1900, La Plata, Buenos Aires, Argentina.
ristvan@frlp.utn.edu.ar; valerial@frlp.utn.edu.ar*

RESUMEN

En Argentina, el fenómeno de la deserción estudiantil universitaria cobra especial interés en las carreras de Ingeniería [1], ya que la evolución de la industria y una creciente vinculación entre empresas y universidades, hacen que estas carreras sean esenciales para consolidar el desarrollo industrial, económico y científico del país.

Como parte del sistema formador de ingenieros y para responder a las demandas propias de la institución, la UTN FRLP se encuentra actualmente desarrollando un sistema de detección temprana de posibles desertores y gestión tutorial, a fin de elevar la tasa de graduación en cada una de las especialidades que brinda.

El objetivo del presente trabajo es identificar los diferentes factores que influyen en la permanencia y abandono de los aspirantes e ingresantes, conformando un sistema de indicadores de detección temprana de problemáticas del aprendizaje y dificultades en el trayecto de formación académica.

La investigación explora la capacidad de la técnica Minería de Datos como una alternativa útil para encontrar información derivada a partir de la detección de patrones de atributos individuales.

Palabras clave: Deserción estudiantil, Educación Superior, Permanencia académica, Retención estudiantil.

CONTEXTO

La línea de investigación presente se encuentra inserta en el Proyecto de Investigación y Desarrollo: “Estudio Sistemático de Deserción Estudiantil Universitaria”, actualmente en etapa de programación, dirigida por un grupo de docentes investigadores del Laboratorio de Ingeniería en Sistemas de Información (LINES) y del Departamento de Ingeniería en Sistemas de Información (DISI) de la Universidad Tecnológica Nacional Facultad Regional La Plata.

El proyecto se encuentra homologado y financiado por la Universidad Tecnológica Nacional, UTN.

1. INTRODUCCIÓN

Himmel define la deserción como “el abandono prematuro de un programa de estudios antes de alcanzar el título o grado, y considera un tiempo suficientemente largo como para descartar la posibilidad de que el estudiante se reincorpore” [2].

A partir de la década de 1970 se desarrollaron los primeros enfoques teóricos, que explican la deserción mediante el análisis de factores individuales (Fishbein y Ajzen [3]), sociológicos (Spady, 1970 [4]), económicos (Cabrera, Castañeda, Nora y Hegnstler 1992 [5]), organizacionales y de interacción (Tinto, 1975 [6]).

Estos modelos iniciales sirvieron de base para investigaciones posteriores, incorporándose en los últimos años enfoques

integrados, que recogen los anteriores en aproximaciones que abarcan a más de una de las líneas enunciadas [7].

Tinto en el año 2010 en una revisión de su modelo, reconoce cuatro condiciones institucionales asociadas con la retención estudiantil: expectativas, apoyo, retroalimentación y participación [8]. En este trabajo enfatiza la necesidad de utilizar varias formas de evaluación y retroalimentación inmediata y sistemática al estudiante, que garanticen una “alerta temprana” para proveer al estudiante el apoyo necesario. Esta última idea es tomada como parte fundamental del presente proyecto.

Situación de Argentina en Latinoamérica y en el mundo

El Centro de Estudios de Educación Argentina (CEA), analizó la evolución de la matrícula y la graduación universitaria entre 2003 y 2012 sobre la base del Anuario de Estadísticas Universitarias, del Departamento de Información Universitaria, la Unesco y la Organización para la Cooperación y el Desarrollo Económicos (OCDE) e identificó que Argentina tiene mayor porcentaje de estudiantes universitarios que Brasil y Chile, teniendo en cuenta la población de cada país, pero el alto nivel de deserción estudiantil determina que proporcionalmente existan solamente 17 estudiantes por cada graduado para marzo de 2015, mientras que en Chile son 8 de cada 4, y 6 de cada 7 en Brasil respectivamente [9].

De esta manera, Argentina se caracteriza por ser uno de los países con menor graduación en proporción al total de estudiantes universitarios en Latinoamérica.

No obstante esta realidad, en 2016 Argentina fue identificada dentro del grupo TACTICS, por el Centro de Educación Superior Global en el Instituto de Educación de la UCL, Times Higher Education del Foro Económico

Mundial. El cual luego de haber analizado indicadores económicos y académicos, identificó a los siete países que estarían en condiciones de convertirse en estrellas en ascenso en Educación Superior, denominadas TACTICS (Tailandia, Argentina, Chile, Turquía, Irán, Colombia y Serbia) y que podrían superar a las naciones del BRICS (Brasil, Rusia, India, China, y Sudáfrica), países identificados tradicionalmente como estrellas emergentes mundiales.

En los países del TACTICS, el PIB es inferior a USD 15.000 por persona, pero al menos la mitad de la población joven está matriculada en Educación Superior y ha crecido un 5% o más entre 2010 y 2014. La cantidad de investigaciones científicas publicadas creció a más de 3.000 por año; y tienen, como mínimo, una universidad rankeada dentro de las clasificaciones de las universidades del mundo de Times Higher Education [10].

A nivel institucional

A nivel institucional, la Universidad Tecnológica Nacional, Facultad Regional La Plata en el año 2016 analizó la relación aspirante / egresado e ingresante / egresado en el período 2010-2015, lo que permitió identificar que 10.48% de los aspirantes se gradúan y que sólo el 15.05% de los ingresantes alcanza con éxito la finalización de su carrera [11].

En la misma línea y focalizando la atención en la relación aspirante / ingresante en el mismo período, cabe mencionar que el 30% de los aspirantes no llegaron a ser ingresantes.

Para hacer frente a este fenómeno, el equipo de investigación se encuentra trabajando en el desarrollo de un sistema informático de detección temprana de posibles desertores; a fin de elevar la tasa de graduación en cada una de las carreras que brinda la UTN Regional La Plata.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

La línea de investigación se encuentra inserta en el Proyecto de Investigación y Desarrollo: “Estudio Sistemático de Deserción Estudiantil Universitaria”.

Como parte esencial de ese proyecto se incluye la particularidad de identificar los diferentes factores que influyen en la permanencia y abandono de los aspirantes e ingresantes, conformando un sistema de indicadores de detección temprana de problemáticas del aprendizaje y dificultades en el trayecto de formación de los estudiantes.

Como instrumentos de investigación para alcanzar los objetivos se utilizan cuestionarios estandarizados y entrevistas; así también como la técnica de Minería de Datos, ya que mediante ella es posible encontrar patrones y relaciones entre dos o más variables analizadas de manera automática, Datawarehouse y Análisis de Redes Sociales (ARS) para la determinación de los conjuntos de interacciones dentro del aula y el status sociométrico.

3. RESULTADOS OBTENIDOS / ESPERADOS

El objetivo principal del proyecto es caracterizar el perfil de los aspirantes e ingresantes de la UTN FRLP, para de esa manera posibilitar la detección tempranamente de posibles desertores y facilitar la aplicación de estrategias de intervención ajustadas a la institución y a cada alumno en particular.

Se plantean como objetivos específicos:

- Identificar los diferentes factores que influyen en la deserción de aspirantes e ingresantes en la UTN FRLP.
- Brindar un modelo explicativo de los principales factores causales de la deserción; caracterizando tempranamente a

posibles desertores a fin de tomar las acciones adecuadas para su retención.

- Desarrollar un sistema de indicadores de riesgo de deserción de ingresantes para la institución, como módulo esencial del futuro sistema en desarrollo.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo del proyecto está formado por docentes investigadores del Laboratorio de Ingeniería en Sistemas de Información (LINES) y del Departamento de Ingeniería en Sistemas de Información (DISI) de la Universidad Tecnológica Nacional Facultad Regional La Plata, un investigador de apoyo, un tesista de postgrado y dos alumnos becarios de investigación.

El proyecto anualmente capacita y forma a alumnos becarios que participan y aprenden desarrollando diversas tareas de investigación. Plantea la modalidad de trabajo colaborativo / cooperativo; a fin de promover la construcción del conocimiento no sólo individual sino también, como instancia de búsqueda de formas trabajo en grupo, fomentando valores como la responsabilidad, comunicación, trabajo en equipo y la autoevaluación individual y de compañeros. Constituyéndose, de esta manera, en una etapa de entrenamiento para la futura actividad profesional y de inicio a la investigación.

El proyecto brinda un marco propicio para el desarrollo de las Prácticas Supervisadas (PS) de los estudiantes, necesarias para la obtención del título de grado de la carrera de Ingeniería en Sistemas de Información. Actualmente se encuentran 2 PS en etapa de ejecución y 1 finalizada.

5. BIBLIOGRAFÍA

- [1] CONFEDI (2010). La Formación del Ingeniero para el Desarrollo Sostenible. Aportes del Congreso Mundial Ingeniería 2010. Buenos Aires.
- [2] Erika Himmel (2002); “Modelos de Análisis de la Deserción Estudiantil en la Educación Superior”; Revista Calidad de la Educación, Chile.
- [3] M. Fishbein y I. Ajzen (1975). Belief, attitude, intention and behavior: An introduction to theory and research. Ed. Addison-Wesley.
- [4] W. Spady (1970). Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange*, 1 (1) 64-85).
- [5] A. Cabrera, M. Castañeda, A. Nora, D. Hengstler (1992). The converge between two theories of college persistence. *Journal of Higher Education*, 63, pg. 143-164.
- [6] V. Tinto (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89-125
- [7] Edward P. St. John, Alberto E Cabrera, Amaury Nora, and Eric H. Asker (2000); Economic Influences on Persistence Reconsidered How Can Finance Research Inform the Reconceptualization of Persistence Models?. 29-47.
- [8] V. Tinto (2010). From theory to action: Exploring the institutional conditions for student retention. (J.C. Smart (ed.), Ed.) (Springer N). Higher education: Handbook of theory and research. doi: 10.1007/978-94-007-2950-6.
- [9] Universia, (2015), Deserción estudiantil: 3 de cada 10 estudiantes se gradúan en argentina, <http://noticias.universia.com.ar/educacion/noticia/2015/03/31/1122497/desercion-estudiantil-3-cada-10-estudiantes-graduan-argentina.html>
- [10] Foro Económico Mundial (2016); These countries could be the world's new education superstars; <https://www.weforum.org/agenda/2016/12/countries-global-higher-education-superstars-tactics/>
- [11] Romina Istvan, Chong Arias Carlos, Antonini Sergio (2016); “Sistema de Indicadores de riesgo de deserción para la UTN-FRLP”, Cytal 2016.
- [12] Mariana Falco, Romina Istvan, Antonini Sergio (2017); University Desertion: Analysis to 2017 admission course in Information Systems Engineering, JAIIO 2017.

UAICASE: ENSEÑANZA DE INGENIERÍA DE SOFTWARE EN ENTORNOS VIRTUALES COLABORATIVOS

Marcelo De Vincenzi, Carlos Neil, Nicolás Battaglia, Roxana Martínez

CAETI – Centro de Altos Estudios en Tecnología Informática
Facultad de Tecnología Informática. UAI – Universidad Abierta Interamericana
{medevincenzi, carlos.neil, nicolas.battaglia, roxana.martinez}@uai.edu.ar

RESUMEN

Hoy en día, existen áreas de conocimiento que requieren de técnicas y herramientas para complementar el proceso de enseñanza y aprendizaje dentro del aula. Es el caso de la Ingeniería de Software (IS), en el que se utilizan tecnologías que ayudan en este proceso, denominadas herramientas CASE (Computer Aided Software Engineering). Además, mediante la inclusión de las TIC en la educación, se generan nuevas oportunidades para el aprendizaje colaborativo, el cual, permite un enfoque basado en la ubicuidad.

En este proyecto, se propone la integración de plataformas tecnológicas para el aprendizaje ubicuo colaborativo en cursos relacionados a la IS, en particular, durante el proceso de modelado de software, donde la solución planteada no reemplaza al modelo presencial, sino que brinda un espacio ubicuo de colaboración para el proceso formativo indicado. Esta integración plantea un nuevo enfoque en la metodología de gestión del trabajo en los diversos equipos colaborativos.

Palabras clave: Trabajo Colaborativo, Enseñanza Virtual Ubicua, Ingeniería de Software, UML.

CONTEXTO

Este proyecto de investigación es desarrollado actualmente en la facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI).

A partir del desarrollo de una herramienta CASE colaborativa, denominada *UAI Case*, los alumnos de la Universidad desde 2° año hasta 5° año de la carrera de Ingeniería en Sistemas Informáticos, trabajan en las siguientes materias relacionadas:

- 1) Metodologías de Desarrollo de Sistemas I y II (2° año).
- 2) Bases de Datos, Trabajo de Diploma y Trabajo de Campo I (3° año).
- 3) Seminario de Aplicación Profesional y Trabajo Final de Ingeniería (5° año).

Con esta herramienta, los alumnos aprenden e interactúan con conceptos que utilizan para aplicar el modelado UML en un proyecto de software en forma colaborativa y evolutiva de forma incremental.

1. INTRODUCCIÓN

La Ingeniería de Software es un área de conocimiento que requiere de software específico, como lo son las herramientas CASE, para facilitar la comprensión y el estudio de modelado UML. Por otro lado, hoy en día, diversas tecnologías se inclinan por el trabajo colaborativo en el ámbito de

la educación, por lo que el Aprendizaje Colaborativo Mediado por Tecnología o CSCL (*Computer Supported Collaborative Learning*) [1] se emplea como un recurso didáctico innovador. Este concepto, “responde a un fenómeno sociocultural actual, donde se define el cómo aprendemos (socialmente) y dónde aprendemos (en red)” [2].

Existen 3 disciplinas en el modelo CSCL: pedagógica, psicológica y tecnológica. Éstas componen un campo de “estudio interdisciplinario donde un número importante de investigadores se han visto motivados a alinear sus esfuerzos en el análisis de los factores intervinientes capaces de potenciar la eficiencia (proceso) y eficacia (resultado) de la aplicación del aprendizaje colaborativo mediado por tecnología en diferentes escenarios educativos” [3]. Por otra parte, es importante tener un proceso de evaluación que permita medir el nivel de aprendizaje realizado hasta el momento, con el fin de identificar los problemas que vayan surgiendo en el proceso de enseñanza.

Las plataformas virtuales requieren ser analizadas y valoradas desde concepciones psico didácticas que permitan su evaluación desde criterios específicamente educativos [4]. Estas plataformas permiten disponer de contenidos en cualquier momento y en cualquier lugar; obligando a disponer de nuevas infraestructuras tecnológicas [5]. Además, la necesidad de una actualización dinámica en forma rápida de los contenidos utilizados, no debe inducir al descuido de conceptos básicos [6].

Como se mencionó anteriormente, los cursos específicos sobre IS, requieren de esfuerzos, no sólo para mejorar la colaboración entre docentes y alumnos, sino también, para realizar evaluaciones y seguimientos de los trabajos universitarios de los estudiantes [7]. En las Instituciones

actuales, los docentes apoyan las nuevas posibilidades de comunicación tales como el soporte para la organización y el dictado de sus materias, generando un gran impacto en la construcción colaborativa del conocimiento [8].

El modelo anteriormente propuesto en este proyecto [9], especifica y vincula las plataformas CSCL con la enseñanza de la Ingeniería de Software en un entorno ubicuo, en particular con el modelado de software por medio de UML.

Como trabajo complementario, se deberán definir métricas para realizar la evaluación empírica de la solución planteada [10] por medio de la implementación del modelo formativo y su uso en cursos homogéneos de la facultad de TI de la UAI. La implementación de estas métricas, facilitará, entre otras cuestiones, la medición de tiempos, la comprensión del sistema a desarrollar y la realización de controles en validaciones de la metodología empleada de colaboración ubicua.

La tarea de evaluar el rendimiento académico de un grupo de estudiantes, comprende, al menos, de dos escenarios distintos. El primer escenario es analizar ~~la~~ el software en sí, y su relación con los alumnos y el docente, con el fin de observar las bondades de la herramienta propuesta, mediante la utilización de conceptos de *empirical software* [11], [12], [13], [14], el segundo escenario está relacionado con la investigación sobre el aprendizaje de herramientas de modelado [15], [16]. Esta última visión consiste en la comparación de nuestra propuesta pedagógica frente a los métodos más tradicionales de la enseñanza de la Ingeniería de Software, como el uso de papel y lápiz para plantear un modelo conceptual o el uso de herramientas CASE comerciales. En este caso, adaptaremos el concepto de ganancia de aprendizaje utilizando el factor de Hake [17], este

factor permite analizar las ganancias de aprendizaje que se han obtenido mediante diferentes procesos de enseñanza y consiste en la aplicación de un mismo instrumento al principio y al final del proceso, bien sea dentro de un mismo grupo o en grupos diferentes. De esta forma nos permitirá comparar cuantitativamente la mejora esperada en el aprendizaje de conocimientos mediante el uso de *UAI Case*, frente a métodos tradicionales.

2. HERRAMIENTAS CASE EN ENTORNOS ACADÉMICOS COLABORATIVOS

Hoy en día, existe un número significativo de trabajos que estudian la evolución de las herramientas CASE con UML, como soporte para la Ingeniería de Software. Éstas, están cada vez más enfocadas en el concepto de la colaboración como herramienta para mejorar los procesos de desarrollo, integrándose en entornos WEB, con todas sus implicancias a nivel tecnología y ubicuidad. Es importante observar que las herramientas actuales de modelado carecen de utilidades destinadas a la enseñanza de la Ingeniería de Software [18].

Por otro lado, el creciente aumento de las tecnologías de las comunicaciones y su próspera integración con los entornos de trabajo colaborativo, abren nuevas puertas a los medios de aprendizaje colaborativos. Este aprendizaje se basa en el desarrollo de estrategias de comprensión y explicación, utilizando debates que sirven para desarrollar las habilidades de comunicación [19].

3. MODELO PROPUESTO: uCASE-CL

A partir de lo antes expresado, surge la importancia de una especificación del

modelo CSCL que permita utilizar una herramienta de dominio específico. El modelo CSCL está planteado con tres dimensiones (Ciencias de la Computación, Psicología y Pedagogía), las cuales resultan afectadas por una cuarta dimensión denominada Ciencias de la Comunicación, en representación de la tecnología subyacente que permite materializar el concepto de ubicuidad.

En [20], los autores explican que los contenidos y las actividades deben ser incorporados utilizando estándares definidos por el modelo CSCL, estableciendo un contenedor de objetos de aprendizaje. Además, proponen una herramienta computacional que integra componentes de *e-learning* para prestar servicios por medio de un modelo de aprendizaje colaborativo de forma ubicua, lo que permitirá a los alumnos acceder a la información desde cualquier parte y lugar mediante dispositivos móviles.

La propuesta del proyecto es definir y extender el modelo, primero para brindar un bloque funcional de evaluación y, segundo, para integrar nuevas herramientas de enseñanza y aprendizaje para aquellas áreas de conocimiento específico, como ser la IS, mediante la propuesta de métricas que evalúen el proceso de aprendizaje colaborativo ubicuo.

En primer lugar, es necesario extender los componentes para los modelos de aprendizaje, ya que los entornos colaborativos de enseñanza comienzan a tener injerencia en el proceso de evaluación. Esto transforma al modelo CSCL tradicional, y le abre nuevas puertas a la evaluación en el entorno colaborativo.

Por otra parte, también es necesario integrar el entorno de aprendizaje con aquellas herramientas propias de la práctica profesional. Estas herramientas deberán contar con recursos colaborativos no solo para la actividad principal de la

temática, sino también para actividades pedagógicas para el aprendizaje, tal como la evaluación colaborativa, dando lugar a la participación activa del docente en el entorno y la posibilidad de explotar al máximo el aprendizaje utilizando la ubicuidad no solo para aprender, sino también para evaluar.

4. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente trabajo está identificado como proyecto de investigación y desarrollo en el CAETI, dentro de la línea de investigación Sociedad del Conocimiento y Tecnologías aplicadas a la educación.

5. RESULTADOS OBTENIDOS/ESPERADOS

5.1. Resultados obtenidos:

- Desarrollo, diseño e implementación de una herramienta CASE colaborativa ubicua multiplataforma, denominada *UAI Case* [21].

- Diseño de una plataforma académica colaborativa multiplataforma para evaluación, seguimiento, interacción y coordinación de proyectos informáticos.

- Definición y especificación de los bloques funcionales necesarios para determinar un proceso específico en la enseñanza y aprendizaje de modelados en la IS.

5.2. Objetivos futuros/Resultados esperados

- Proponer una metodología de desarrollo de sistemas para ser utilizado en las asignaturas de integración curricular de la carrera Ingeniería en Sistemas Informáticos.

- Orientar el proyecto en el ámbito académico (interno y externo) con el fin de dar visibilidad al trabajo desarrollado por

la facultad de TI de la UAI de modo tal que la herramienta CASE pueda ser implementada por otras carreras.

- Obtener un conjunto de métricas que permitan medir el rendimiento de un equipo virtual de trabajo durante el proceso de enseñanza y aprendizaje de la IS.

- Realizar la evaluación empírica del prototipo de la herramienta *UAI Case* por medio de la definición de un método de evaluación basado en métricas. En una primera instancia, evaluaremos a dos grupos de alumnos de la asignatura Metodología de Desarrollo de Sistemas I; en uno de ellos utilizaremos la herramienta *UAI Case*, en el otro (grupo de control) usaremos el formato tradicional de enseñanza aprendizaje. El objetivo es medir la ganancia de aprendizaje que nos permita evaluar en qué medida el uso de la herramienta es una instancia superadora del método tradicional.

6. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado, además de los integrantes que encabezan este trabajo, por los docentes de las asignaturas vinculadas con el proyecto de integración curricular. También participarán 8 alumnos de las diferentes asignaturas de la carrera. Al momento, en base a este trabajo de investigación, surgió una tesis de maestría en tecnología informática aprobada y otra tesis doctoral en proceso.

7. BIBLIOGRAFÍA

1. Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. *Cambridge handbook of the learning sciences*, 2006, 409-426.
2. Zañurtu Correa, Luz M. (2006). Art.: Aprendizaje colaborativo: una nueva

- forma de diálogo interpersonal y en red. Revista digital de educación y nuevas tecnologías. Contexto educativo. Nro 28, Año 5.
3. Guisen, A., Sanz, C. V., & De Giusti, A. E. (2009). Sistemas cscl (computer supported collaborative learning) para saac. In XV Congreso Argentino de Ciencias de la Computación.
 4. Pons, J. D. P., Bravo, M. P. C., & Ramírez, T. G. (2016). La enseñanza universitaria apoyada en plataformas virtuales. Cambios en las prácticas docentes: el caso de la Universidad de Sevilla. *Estudios sobre Educación*, 20, 23-48.
 5. Filippi, J. L., Lafuente, G. J., & Bertone, R. A. (2010). Diseño de un ambiente de aprendizaje colaborativo. In V Congreso de Tecnología en Educación y Educación en Tecnología.
 6. Daniele, M., Uva, M., Martelloto, P., & Picco, G. (2010). Aplicación de herramientas CASE a la enseñanza de Ingeniería de Software: Gestión de la Configuración de Software y testing Funcional. In V Congreso de Tecnología en Educación y Educación en Tecnología.
 7. Neil C., De Vincenzi M., Battaglia N., Martínez R. (2016). Herramientas Colaborativas Multiplataforma en la Enseñanza de la Ingeniería de Software. In XVIII Workshop de Investigadores en Ciencias de la Computación.
 8. Oliveros, A., & Piccin, A. M. (2012). Web 2.0 para el ámbito académico: Implicancias en la Ingeniería de Software. In XIV Workshop de Investigadores en Ciencias de la Computación.
 9. Battaglia, N., Neil C., Martínez R., González, D., De Vincenzi M. (2017). Learning of Software Engineering on Collaborative Virtual Environments. In 7th Word Engineering Education Forum (WEEF).
 10. Battaglia, N., Neil, C., De Vincenzi, M., & Martínez, R. (2016). UAI Case: integración de un entorno académico con una herramienta CASE en una plataforma virtual colaborativa. In XI Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2016).
 11. Juristo, N., & Moreno, A. M. (2013). Basics of software engineering experimentation. Springer Science & Business Media.
 12. Sjøberg, D. I., Hannay, J. E., Hansen, O., Kampenes, V. B., Karahasanovic, A., Liborg, N. K., & Rekdal, A. C. (2005). A survey of controlled experiments in software engineering. *IEEE transactions on software engineering*, 31(9), 733-753.
 13. Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). Experimentation in software engineering. Springer Science & Business Media.
 14. Shull, F., Singer, J., & Sjøberg, D. I. (Eds.). (2007). Guide to advanced empirical software engineering. Springer Science & Business Media.
 15. Soler, J., Boada, I., Prados, F., Poch, J., & Fabregat, R. (2010, April). A web-based e-learning tool for UML class diagrams. In Education Engineering (EDUCON), 2010 IEEE (pp. 973-979). IEEE.
 16. Siau, K., & Loo, P. P. (2006). Identifying difficulties in learning UML. *Information Systems Management*, 23(3), 43-51.
 17. Hake, R. R. (1998). Interactive-Engagement vs. Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses.
 18. Del Ben, E., & Vinjoy, M. (2011). Tecnología para la enseñanza de bases de datos: Un entorno de desarrollo

- enfocado a la enseñanza del diseño de bases de datos. In XIII Workshop de Investigadores en Ciencias de la Computación.
19. Lavigne, G., Ovando, M. P. V., Sandoval, J. O., & Salas, L. M. (1970). Exploración preliminar del aprendizaje colaborativo dentro un entorno virtual/Preliminary study of collaborative learning in a virtual environment. *Revista Actualidades Investigativas en Educación*, 12(3).
 20. Coto, M., Collazos, C. A., & Rivera, S. M. (2016). Modelo Colaborativo y Ubicuo para apoyar los procesos de enseñanza-aprendizaje a nivel Iberoamericano. *Revista de Educación a Distancia*, (48).
 21. Neil C., De Vincenzi M., Battaglia N., Martínez R. (2016). Herramientas Colaborativas Multiplataforma en la Enseñanza de la Ingeniería de Software. In XVIII Workshop de Investigadores en Ciencias de la Computación.

Procedimientos y herramientas para la mejora de indicadores académicos en facultades de ingeniería.

Gustavo Illescas, Martín Santiago¹, Mariano Martínez

Instituto de Investigación en Tecnología Informática Avanzada, Grupo en Informática de Gestión - Centro Asociado CIC. Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires. Paraje Arroyo Seco, Tandil (7000), Argentina. +54 249 4385680

illescas@exa.unicen.edu.ar, msantiag@exa.unicen.edu.ar, mmartinez@slab.exa.unicen.edu.ar

RESUMEN

El objetivo del proyecto presentado es el diseño y desarrollo de procedimientos y herramientas de software que asistan a la mejora de indicadores académicos en las facultades de ingeniería de la Universidad Nacional del Centro de la Provincia de Buenos Aires en cuanto a los programas de Ingreso, Permanencia y Graduación.

Los principales aportes de este proyecto consisten en la introducción del paradigma de gestión por indicadores en ámbitos jerárquicos de las unidades académicas de la Universidad que dictan carreras de Ingeniería y en despertar interés en otras Facultades.

También ha dado lugar al desarrollo de prototipos altamente funcionales que permiten la implantación práctica de este tipo de tecnologías a partir de la puesta en marcha de dos tesis de grado de alumnos de Ingeniería de Sistemas. Para lograr esto se han sorteado obstáculos y refinado detalles para otorgar a los prototipos las características de un producto final, y más allá de que los tesis cumplan con sus objetivos académicos, se espera que estas herramientas puedan también propagarse a otros espacios públicos y/o privados.

Palabras clave:

Indicadores académicos. Mejora continua. Control de gestión. Cuadro de mando. Tablero de control.

CONTEXTO

La propuesta emerge como parte de las actividades llevadas adelante en el marco de la implementación en el ámbito de la Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA) del Plan Estratégico de Formación de Ingenieros (PEFI), impulsado por la Secretaría de Políticas Universitarias, y que tiene como objetivo incrementar la cantidad de graduados en ingeniería².

Para el desarrollo de las actividades se cuenta con integrantes del grupo de investigación en Informática de Gestión (Instituto de Investigación en Tecnología Informática Avanzada-INTIA) y alumnos de grado y posgrado de la Facultad de Ciencias Exactas (FCEX) de la UNCPBA, como así también con la colaboración de los equipos de gestión y de las áreas TIC de las Facultades de Agronomía (FAA), Ingeniería (FIO) y FCEX.

Las tres facultades involucradas participan como adoptantes de la solución mediante la implementación de las herramientas en servidores propios.

1. INTRODUCCIÓN

La propuesta se centra en la metodología conocida como “Tableros de control”. En los

¹ Docente-Investigador del Instituto de Física Arroyo Seco (IFAS), Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires. Vice decano al inicio del proyecto.

² <http://pefi.siu.edu.ar/>

tableros se definen y agrupan indicadores por temática, área de la organización o por interés del observador (Ballvé 2000), donde la imagen usualmente utilizada para comprender el concepto es el tablero de un avión.

Al comienzo del desarrollo del proyecto se identificaron requerimientos a partir de la pregunta: “*Qué deberíamos esperar de una herramienta para la Gestión por indicadores*”. Se acordó en que la herramienta debería ser web y desarrollada con software libre. Además, debería permitir crear sin limitación nuevos indicadores que sean:

- parametrizables
- accesibles por múltiples usuarios
- actualizables de manera automática y fácilmente accesibles
- agrupables en diferentes tableros por áreas, temáticas, interés del usuario, etc.
- visualizables en líneas de tiempo
- exportables a planillas de cálculo, páginas web públicas o reportes estructurados

Por otra parte, se exploró sobre “*Qué software disponible podemos utilizar en la Gestión por indicadores*”. Por este motivo se ha realizado un relevamiento de herramientas como parte del trabajo de investigación de dos tesis de ingeniería de sistemas y una tesis doctoral a partir de búsquedas en internet. Luego de realizar pruebas sobre algunas de estas herramientas, se observó en general: escasa documentación, difícil configuración y puesta en marcha, interfaz poco amigable e imposibilidad de acceso a la base de datos (BD) (Illescas 2014).

Dado este escenario se decidió proponer el tema como parte de una tesis de grado de ingeniería de sistemas, basados en las premisas mencionadas. La tesis se denomina “*Herramienta para dar soporte a la gestión de indicadores*”.

Debido a que esta herramienta tiene su propia base de datos, como parte de un proyecto integral se ha propuesto el desarrollo de una tesis en paralelo a la anterior para la construcción de un software que permita conectar las bases de datos de los sistemas

operacionales (o también llamadas BD Externas- BDExt) con la base de datos que contiene los indicadores (BDI), tal que estos puedan ser actualizados en forma sistematizada de manera automática o semi-automática. Esta segunda tesis lleva el nombre de “*Integrador de fuentes de datos aplicado a la gestión por indicadores*”.

Para el caso concreto de aplicación y debido a la necesidad de intervenir en casos de estudio referidos a la gestión académica, se trabajó con BDs del sistema SIU-Guaraní³.

Es importante señalar que las tesis de grado tienen la finalidad de demostrar que se puede cumplir con un objetivo de diseño hasta la fase de prototipación. En ningún caso se llega a la fase de producto ya que excede a la dedicación del alumno. Para poder pasar de una fase a otra, las tres unidades académicas tomaron la iniciativa de destinar recursos para sostener el trabajo de los tesisistas comprometidos en el desarrollo de funcionalidades específicas que potenciarán la capacidad de los prototipos para volverlos utilizables al nivel de un producto final.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación del tema que se está investigando integran un conjunto de herramientas y soluciones de asistencia a la toma de decisiones basadas en indicadores y, como se mencionó anteriormente, se corresponden con dos tesis de grado presentadas para la carrera de grado en Ingeniería de Sistemas (FCEX-UNCPBA). Las referencias citadas en cada línea se corresponden con las principales fuentes de información para el tratamiento de cada de dichas tesis de grado.

Línea 1. Herramienta para dar soporte a la gestión por indicadores. (AENOR 2003; Ballvé 2000; Illescas et al. 2014; Kaplan y Norton 2000; Olve et al 2000).

Línea 2. Integrador de fuentes de datos aplicado a la gestión por indicadores.

³ Desarrollado por el consorcio SIU:
<https://www.siu.edu.ar/>

(Berenson et al 1996, Del Sordo et al 2012; Illescas 2014; Mora-Soto J. 2011).

Adicionalmente se encuentra en elaboración un plan de trabajo para una nueva tesis de grado que complementa las anteriores: *Línea 3. Herramienta para brindar soporte a la gestión por mapas estratégicos* (Barone et al 2011; Illescas 2014; Kaplan y Norton 2000 y 2004; Ruskov et al 2008).

Los alumnos que están elaborando las tesis son mencionados en la estructura del equipo de trabajo.

Diseño de indicadores

Como primer paso se analizó la documentación referida a los informes provistos por el Sistema de estadísticas de alumnos SIU-Araucano⁴ con el fin de identificar los indicadores usuales en la gestión académica. Con posterioridad y teniendo esto último en mente, se realizaron diversas reuniones a las que fueron convocados personas de distintas secretarías y departamentos para convenir la necesidad de analizar la información provista por los sistemas SIU y la posibilidad de transformarlos en indicadores de gestión.

Como conclusión de esta actividad se puede mencionar que se seleccionaron como prueba piloto los indicadores básicos de ingreso y graduación de cada carrera (aspirantes, inscriptos, reinscritos, egresados). Por otro lado, se definieron indicadores de *eficacia* (nivel de logro), *eficiencia* (nivel de esfuerzo requerido para el nivel de logro alcanzado), *eficiencia y eficacia por cohorte*, *desgranamiento*, entre otros.

3. RESULTADOS OBTENIDOS

Se detallan a continuación las actividades que ya han sido realizadas como avances fundamentales para llevar a cabo el proyecto propuesto.

Línea 1. Herramienta para dar soporte a la gestión de indicadores (Fig. 1 y 2).

Etapa 1: Desarrollo de funcionalidades de creación de tableros, indicadores y usuarios. Configuración de umbrales. Registro de umbrales históricos para su representación gráfica. Indicar si el indicador es ascendente o descendente.

Etapa 2: Desarrollo de herramientas de representación gráfica (versión preliminar) y descarga de datos. Dar la posibilidad de cambiar periodos para mostrar en gráficos de históricos.

Etapa 3: Reportes (versión preliminar). Importación, agregado y eliminación de datos.

Etapa 4: *Login* y manejo de sesiones. Representación gráfica y reportes (versión final). Testeo general e instalación en el servidor central.



Figura 1. Indicadores definidos en el tablero "Estudiantes Exactas". Los datos mostrados son simulados.



Figura 2. Gráficos representativos del Indicador "Nuevos Inscriptos sistemas". Los datos mostrados son simulados.

Línea 2. Integrador de fuentes de datos aplicado a la gestión por indicadores (Fig. 3).

Etapa 1: Diseño de la solución y selección de herramientas de desarrollo. Instalación y prueba de las herramientas de desarrollo

⁴ <http://araucano.siu.edu.ar/>

seleccionadas. Escritura del diseño de la solución.

Etapas 2: Desarrollo del *frame* y conectividad BD. Configuración de la BDI para acceder a datos a ser actualizados.

Etapas 3: Generación de la estructura para hacer las consultas a BDExt tales como los sistemas SIU y registro automatizado de la actualización de datos a la BDI.

Etapas 4: Tablero de ejecución según calendario: Generar un calendario donde se pueda visualizar el estado de las actualizaciones, que refleje si las mismas se ejecutaron con éxito. Testeo general e instalación en el servidor central.



Figura 3. Ejemplo de una consulta a la BD Guaraní en la consola del editor de código.

Línea 3. Herramienta para brindar soporte a la gestión por mapas estratégicos

Elaboración del plan de trabajo para ser presentado como propuesta de tesis de grado en la carrera de Ingeniería de Sistemas (FCEX).

Al momento de esta publicación las tesis de grado que abarcan los ejes principales del tema que se está investigando (líneas 1 y 2) se encuentran es la etapa final de escritura del informe final.

Se han instalado los prototipos en los tres servidores de las facultades involucradas y se han realizado pruebas satisfactorias. Se dictaron capacitaciones a los usuarios finales (personal del equipo de gestión), utilizando los prototipos.

Adicionalmente se resolvieron diversas dificultades que devienen de la integración con sistemas reales en producción. Por este

motivo, se puede considerar que los prototipos son viables como productos finales.

Como parte de los trabajos a futuro resta la puesta en marcha definitiva por parte de las facultades intervinientes lo que conlleva la definición de tableros e indicadores a utilizar por parte de los equipos de gestión. Durante esta actividad, el equipo de desarrollo acompaña en el proceso atendiendo posibles errores y cuestiones de mantenimiento.

Por último, se espera avanzar con éxito en el desarrollo de la tesis mencionada en la línea 3.

4. FORMACIÓN DE RECURSOS HUMANOS

La estructura del equipo de trabajo se muestra en la siguiente tabla:

Apellido y nombre	Título	Cargo	Funciones
Illescas Gustavo	Dr.	Prof. UNCPBA	Coordinador general
Santiago Martín	Dr.	Prof. UNCPBA	Analista de requerimientos
Mariano Martínez	Ing. Tesis en curso	Alumno de posgrado	Coordinador técnico
Etchepare Federico	Tesis en curso	Alumno de grado	Integrante
Servat Agustín	Tesis en curso	Alumno de grado	Integrante
Figini Iris	Tesis en curso	Alumno de grado	Integrante
Solanilla Baltasar	Tesis en curso	Alumno de grado	Integrante
Deccechis Juan	Tesis en curso	Alumno de grado	Integrante

5. BIBLIOGRAFIA

AENOR (2003). Guía para la implantación de indicadores. Norma Española UNE66175. *Asociación Española de Normalización y Certificación. Sistemas de gestión de calidad.* Madrid, España.

Ballvé A. (2000). Tablero de control. Ed. Macchi. Buenos Aires.

Barone D.; Jiang L.; Amyot D.; Mylopoulos J. (2011). "Composite indicators for business intelligence". Jeusfeld M., Delcambre L., and Ling T.W. (Eds.) © Springer-Verlag Berlin Heidelberg.

Berenson M.; Levine D. (1996) *Estadística básica en administración, conceptos y aplicaciones*. Sexta edición. Prentice Hall Hispanoamericana, S. A. México.

Del Sordo C.; Orelli R.; Padovani E.; Gardini S. (2012). *Assessing Global Performance in Universities: An Application of Balanced Scorecard*. Elsevier Ltd.

Illescas G. (2014). *Aplicación de métodos matemáticos en el control de gestión por indicadores*. Tesis Doctoral en Matemática Computacional e Industrial, (FCEX-UNCPBA). Tandil. En repositorio digital de la biblioteca central.

Illescas G.; Sánchez-Segura M.; Canziani G.; Xodo D. (2016). *XVIII Workshop de Investigadores en Ciencias de la Computación*. Universidad Nacional de Entre Ríos.

Kaplan R.; Norton D. (2000). *Cuadro de mando integral*. 2da edición. Ediciones GESTION 2000.

Kaplan R.; Norton D. (2004). *Mapas estratégicos*. Ediciones GESTION 2000.

Mora-Soto J. (2011). "Marco metodológico y tecnológico para la gestión del conocimiento organizativo que de soporte al despliegue de buenas prácticas de ingeniería de software". Tesis Doctoral. Escuela Politécnica Superior, Universidad Carlos III, Madrid. Dirección: Dra. Sánchez-Segura M.

Olve N.; Roy J.; Wetter M. (2000). *Implantando y gestionando el cuadro de mando integral*. Ed. Gestión 2000. Barcelona.

Ruskov P.; Todorova Y. (2008). Learning and growth strategy metrics. *International Conference on Computer Systems and Technologies-CompSysTech'08* ©ACM.

SINERGIA1. "GUÍA PARA ELABORACIÓN DE INDICADORES". Departamento Nacional de Planeación. Sistema Nacional de Evaluación de Resultados de la Gestión Pública (SINERGIA). Colombia. Grupo Asesor de la Gestión de Programas y Proyectos de Inversión. Pública (GAPI).

Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos

Sanz Cecilia^{1,5}, Madoz Cristina¹, Gorga Gladys¹, Gonzalez Alejandro¹, Zangara Alejandra¹, Depetris Beatriz², Iglesias Luciano¹, Ibáñez Eduardo¹, Artola Verónica^{1,4}, Violini Lucía^{1,3}, Salazar Mesia Natalí^{1,3}, Sanchez Mariano¹

¹Instituto de Investigación en Informática LIDI (III-LIDI), Centro Asociado CIC,
Facultad de Informática – Universidad Nacional de La Plata

²Universidad Nacional de Tierra del Fuego

³Becaria TIPO A UNLP

⁴Becaria Doctoral CONICET

⁵Investigador Asociado de CIC

{csanz, cmadoz, ggorga, agonzalez, li, eibanez, vartola, lviolini, nsalazar, msanchez}@lidi.info.unlp.edu.ar,
alejandra.zangara@gmail.com, depetrisb@gmail.com

RESUMEN

El presente trabajo describe las principales líneas de investigación y resultados obtenidos en el subproyecto “Escenarios educativos mediados por Tecnologías de la Información y la Comunicación (TIC)”. Este proyecto continúa en 2018 con un nuevo subproyecto denominado “Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos”, enmarcado en el proyecto “Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso”.

En particular, este trabajo comparte los resultados y avances realizados durante 2017 e inicios de 2018.

El subproyecto tiene como objetivo general investigar y realizar aportes, al área disciplinar que integra, de las posibilidades de las Ciencias de la Computación en la Educación. Se investiga y se desarrollan aplicaciones, herramientas y metodologías que permiten poner en diálogo a las tecnologías digitales con los procesos de enseñar y aprender.

Se cuenta con un equipo de trabajo interdisciplinario, en el que participan investigadores formados, en formación, y becarios.

Palabras clave: Trabajo y aprendizaje colaborativo, Entornos digitales, Materiales educativos digitales, Objetos de aprendizaje.

CONTEXTO

Este subproyecto forma parte del proyecto “Tecnologías para Sistemas de Software Distribuidos. Calidad En Sistemas y Procesos. Escenarios Educativos Mediados Por TIC (período 2014-2017)”, perteneciente al Instituto de

Investigación en Informática LIDI, de la Facultad de Informática de la Universidad Nacional de La Plata y acreditado por el Ministerio de Educación de la Nación.

1. INTRODUCCION

La creciente disponibilidad y acceso a dispositivos móviles, a redes inalámbricas, a sensores, sumado al uso del GPS y otras tecnologías que hoy en día se encuentran al alcance de las personas, han marcado la profundización de nuevos entramados entre mundo físico y virtual y la puesta en práctica de los principios de la computación ubicua. Al mismo tiempo, las posibilidades de la web y sus desarrollos continúan siendo tema de investigación y actualidad. Estos escenarios son reconocidos como híbridos, en ellos se aumenta computacionalmente a diferentes objetos del entorno y es posible acceder a la información digital, ya no sólo desde una PC de escritorio, sino desde dispositivos con variadas características. Sin duda, estos cambios que se vienen gestando propician la necesidad de innovaciones para el escenario educativo y hacen necesario fortalecer la investigación sobre cómo las personas, y en particular los docentes y los alumnos, se relacionan con estas tecnologías, para su apropiación en el marco de los procesos de enseñar y aprender. Como menciona Cabero Almenara (2015) el docente cuenta con un “verdadero ecosistema digital que nos debe llevar a replantearnos algunas de las ideas que hemos manejado sobre la incorporación de las TIC a los contextos de formación”. Son varios los autores y los reportes que indican que la colaboración, la generación de habilidades de autorregulación de los aprendizajes, y las destrezas para la utilización de las TIC, son y serán factores claves en la educación actual y futura (Khun, 2017), (Griffin et al., 2012). Es decir, son las personas, y no la tecnología,

quienes aportarán a la generación de reales innovaciones a partir de la apropiación de las tecnologías disponibles.

En este sentido, el subproyecto que aquí se presenta se focaliza, a través de sus líneas de investigación, en esos aspectos ya que:

- Indaga sobre los procesos, herramientas y metodologías para la colaboración en entornos educativos mediados por TIC;
- Investiga sobre diseño y producción de materiales educativos digitales que aprovechen los lenguajes multimediales e hipermediales, que se enfocan en diferentes grados de granularidad acorde al escenario de trabajo y las necesidades del contexto;
- Estudia y realiza experiencias vinculadas a las posibilidades de los entornos digitales. Se analizan y comparan diferentes tipos de entornos y se pone el énfasis en su impacto para la mejora de los procesos de enseñar y aprender; la participación activa del estudiante y el docente; la integración de herramientas cognitivas, metacognitivas y afectivas para el aprendizaje autónomo del estudiante.

En síntesis se abordan líneas de investigación actuales en el campo de la Ciencia de la Computación y se recapitaliza la evolución de las tecnologías asociadas a esta disciplina, en función de las necesidades del escenario educativo, por lo que no sólo se realizan desarrollos (software y hardware), sino que se realiza investigación sobre cómo generar estrategias y metodologías, que impacten en su apropiación por parte de los actores de la educación para favorecer los procesos de enseñar y aprender.

2. LINEAS DE INVESTIGACION / DESARROLLO

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- Entornos digitales: campus virtuales, entornos virtuales de enseñanza y aprendizaje, entornos virtuales 3D, entornos simulados y laboratorios, y redes sociales, entre otros. Proyectos que los incorporan, metodologías, desarrollos, evaluación de su calidad y experiencias.
- Materiales educativos digitales. Metodologías para su diseño y producción. Objetos de aprendizaje. Multimedia e hipermedia en escenarios educativos. Nuevos entramados de medios, soportes y lenguajes.
- Integración de TIC en procesos educativos. Análisis de las actitudes y percepciones de los

docentes. Hibridación de las modalidades educativas.

- Trabajo colaborativo mediado por TICs. Autorregulación y capacidades metacognitivas como factores claves para su desarrollo. Conceptualización, análisis y desarrollo de software y metodologías. Actividades colaborativas aprovechando dispositivos móviles.

3. RESULTADOS OBTENIDOS

Se presentan aquí los principales resultados que se han obtenido en el subproyecto, que son el foco de este trabajo y corresponden al período 2017 - inicios de 2018.

Entornos digitales: campus virtuales, entornos virtuales de enseñanza y aprendizaje (EVEAs), entornos virtuales 3D, entornos simulados y laboratorios, redes sociales.

Esta línea aborda el estudio sobre las posibilidades de diferentes tipos de entornos digitales en procesos educativos. Se analizan las divergencias, convergencias y tensiones entre los entornos sociales (redes sociales, *microblogging*) y entornos institucionales. Mientras que los primeros impactaron en la interacción social entre las personas, en el aprendizaje informal y en la investigación de estos como parte del desarrollo de Entornos Personales de Aprendizaje (EPA o PLE: *Personal Learning Environments*)¹ (Castañeda, Dabbagh & Torres Kompen, 2017), los segundos han formado parte de las propuestas institucionales para extender o invertir el aula, mediar procesos educativos e innovar en la enseñanza y el aprendizaje. Tal como afirman (Dabbagh & Fake, 2017): “Mientras que la mayoría de instituciones de enseñanza superior todavía dependen principalmente del LMS [*Learning management system*] para impartir enseñanza y facilitar el aprendizaje, el debate ha avanzado hacia la integración de los medios sociales en un LMS o, más importante todavía, en cómo las instituciones pueden usar redes sociales existentes para apoyar la creación de PLE por parte de sus alumnos”.

En esta línea en el marco del proyecto, durante el 2017, se ha avanzado en el diseño y evolución del

¹ Según (Dabbagh & Fake, 2017): un EPA es “un ecosistema formado libremente que consiste en cualquier conjunto de canales de comunicación, recursos en la nube, aplicaciones web y usuarios de redes sociales”, “y un enfoque pedagógico potencialmente prometedor para la integración del aprendizaje formal e informal y la promoción del aprendizaje autorregulado”

entorno de enseñanza y aprendizaje IDEAS², al que se le ha realizado un estudio sobre accesibilidad, a partir de un trabajo de grado (Galimberti, Sanchez, Gorga & Sanz, 2016). Para ello, se analizaron los resultados de diferentes pruebas de accesibilidad y se trabajó específicamente con personas con discapacidad visual. Además, se integró una herramienta prototipo para la escritura de mensajes de comunicación aumentativa dentro del editor de textos de IDEAS (Santillán, Gorga & Sanz, 2017). Se han diseñado nuevas funcionalidades para facilitar la participación en el foro de debate de IDEAS, por ejemplo, se incorporó la visualización de un árbol que presenta el mapa de hilos de comunicación para facilitar el seguimiento del debate y la interacción entre los participantes (Iglesias & Archuby, 2017). En este sentido se avanzó sobre trabajos previos del equipo de investigación (Sanz, Zangara & Dieser, 2016). También se ha iniciado la creación e integración de una herramienta de comunicación sincrónica para reuniones, denominada *Infomeeting* y que forma parte de una tesina de grado (Digiani, Sanz & Gorga, 2017). Al mismo tiempo, se integró la herramienta de autoevaluación de IDEAS a un juego para dispositivos móviles que permite presentar las preguntas de una evaluación con una serie de desafíos en el marco de una aventura pirata (Archuby, Sanz & Pesado, 2017).

En relación a los entornos sociales, se llevó adelante una experiencia de trabajo colaborativo³, que integró el uso de *Twitter*, un sitio web armado con *Google Sites* y el EVEA IDEAS. De esta manera, se modelizó una estrategia para la integración de diversos entornos: sociales e institucionales. Además, los alumnos eligieron sus propias herramientas de comunicación y producción para llevar adelante la tarea, por lo que generaron su propio ecosistema digital, al decir de Dabbagh et al. (2017).

Se aprobó una tesis de maestría dirigida por integrantes del equipo sobre el diseño de un curso en un EVEA para la enseñanza de inglés (Simón, Zangara & González, 2017).

Otras líneas trabajadas en torno a este eje temático, ha sido el de la autorregulación del aprendizaje en procesos educativos mediados por EVEAS, y entornos digitales en general, a través de una tesis doctoral (Fierro, Sanz & Zangara, 2017) y un trabajo de especialización (Dieser, Sanz & Zangara, 2017).

² IDEAS es una evolución del EVEA WebUNLP y se encuentra en uso en la Facultad de Informática en diferentes cursos de grado y postgrado.

³ Esta experiencia fue desarrollada en el marco del Seminario de Educación a Distancia de la Maestría en TIAE de la Fac. de Informática de la UNLP.

También, se continúa investigando sobre la integración de estos entornos con repositorios de Objetos de Aprendizaje para la implementación de cursos abiertos masivos on-line (MOOC). Además, se analizan las posibilidades de los MOOC y se está avanzando en un trabajo de Especialización sobre esta temática (Pietroboni & Sanz, 2017).

Respecto de la investigación sobre mundos virtuales 3D y entornos inmersivos, se continuó con una tesis doctoral que indaga sobre las posibilidades de este tipo de entornos para personas con hipoacusia (Fachal, Sanz & Abásolo, 2016).

Diseño, producción y evaluación de Materiales educativos digitales. Objetos de aprendizaje y sus Repositorios.

En 2017 se consolidaron los avances que se gestan desde 2015 en relación a la temática de Objetos de Aprendizaje (OA). Se aplicó la metodología de diseño de OA (CROA), elaborada durante 2014/2015 (Sanz, Barranquero & Moralejo, 2016) en diferentes espacios en el marco de la Facultad de Informática. Se trabajó en conjunto con la UNNOBA en el diseño y desarrollo de un objeto de aprendizaje que introduce el concepto de autorregulación. Este objeto se puso a disposición de los estudiantes ingresantes 2018 a la Facultad de Informática de la UNLP⁴ y se evaluará luego su impacto en alumnos de primer año.

Al mismo tiempo, se desarrollaron otros objetos de aprendizaje que estratégicamente buscan complementar el desarrollo de temas de las carreras en Informática: uno orientado a comprender el *merge* de listas ordenadas (Artola & Sanz, 2017), otro sobre la inserción de un nuevo valor en un vector ordenado (Violini & Sanz, 2017), otro sobre compuertas lógicas (Estrebou, Salazar & Sanz, 2017) y uno orientado a comprender la modularización (Archuby, 2017).

Se continuo con la investigación sobre sistemas recomendadores de objetos de aprendizaje y se presentaron algunas líneas de investigación referidas a este tema en el curso de doctorado de Diseño y Producción de OA de la Facultad de Informática (Sanz & Barranquero, 2017) y en un tutorial que se dictó en el marco de LACLO 2017 (Sanz, Barranquero & Astudillo, 2017).

También se ha avanzado en el desarrollo de una tesis doctoral que propone la investigación y desarrollo de *frameworks* para la creación de OA. Durante el 2017 se inició el diseño del primer prototipo del *framework MarCOA* y se hizo una primera evaluación con la opinión de expertos (Violini, Sanz & Pesado, 2017).

⁴ <http://163.10.22.82/OAS/autorregulacion/index.html>

Se finalizó una tesis de maestría sobre materiales hipermediales en la enseñanza de la Matemática. La tesis analiza los marcos teóricos que fundamentan la integración de este tipo de materiales en la enseñanza de la Matemática (Del Río, Bucari & Sanz, 2017).

Se finalizó una tesina de grado cuyo propósito es incorporar al entorno CMRE (ya existente) conceptos de Concurrencia y Paralelismo relacionados con consumo de energía, temperatura y balance de carga de modo de facilitar su aprendizaje a alumnos de los primeros años de Informática (Castro, De Giusti & Gorga 2017).

Se continúa con la investigación sobre la utilización de materiales educativos y actividades de Realidad Aumentada (RA) en procesos educativos. Se avanza en un trabajo de especialización con una becaria del equipo de investigación (Salazar, Sanz & Gorga, 2018), y se participa en diferentes proyectos con distintas universidades del país y del exterior en este tema. Por ejemplo, se dirige un proyecto de la Universidad de Río Negro, en el que se ha investigado sobre la evaluación de materiales que incluyen RA (Lovos, Gibelli & Sanz, 2017).

Trabajo colaborativo mediado por TICs.

En esta línea se abordan metodologías específicas para llevar adelante procesos colaborativos mediados por TICs. Se analizan e implementan herramientas para mediar estos procesos. Se ha trabajado con una metodología de seguimiento de la colaboración, a partir de la definición de indicadores, que surgen de un estudio bibliográfico en profundidad y se incluye una estrategia de *mirroring*, para reflejar las interacciones del grupo participante (Zangara & Sanz, 2017). Se ha finalizado, y entregado para su evaluación, una tesis doctoral vinculada a estos temas de un miembro del proyecto y dirigida por otro miembro del mismo.

Se continúa trabajando con un grupo de la Universidad de Santiago del Estero en el que se llevan adelante experiencias colaborativas con dispositivos móviles (Herrera, Palavecino & Sanz, 2017).

Se profundizó en la aplicación de la metodología de seguimiento del trabajo en foro presentada en 2016 (Sanz, Zangara & Dieser, 2016).

Al mismo tiempo en el proyecto se abordan temas transversales como la integración de TIC en procesos educativos, las dimensiones de análisis de estos procesos, y los actores involucrados, revisando factores claves como la capacitación y las actitudes. En estos temas se está desarrollando una tesis de maestría que propone un modelo de análisis de la integración de TIC en una institución de

educación superior, y en la que se lleva adelante un estudio de caso (Assinato, Sanz & Gorga, 2017).

En cuanto a los proyectos vinculados con la temática y los acuerdos de cooperación, el III- LIDI participa en los siguientes:

- Se ha firmado un acuerdo de colaboración en estos temas con la Universidad de Zaragoza y se trabaja en forma conjunta. Durante el 2017 se participó en un proyecto de innovación docente de la Universidad de Zaragoza (PIIDUZ) y se continúa en 2018. Proyecto: Diseño e implantación de una estrategia de enseñanza-aprendizaje activa centrada en el estudiante y basada en aula invertida y la herramienta M-eroDes. PIIDUZ_16_047.
- Se participa en la Red constituida por universidades de Iberoamérica en el marco del programa "Pablo Neruda" dentro del Espacio Iberoamericano del Conocimiento (EIC) y de la Organización de los Estados Iberoamericanos (OEI) orientada a la movilidad de estudiantes y docentes de doctorado.
- RedAUTI: Red temática en Aplicaciones y Usabilidad de la Televisión Digital Interactiva. En este ámbito se analizan materiales educativos para la TVDI.
- REFORTICCA: Recursos para el Empoderamiento de FORMadores en TIC, Ciencias y Ambiente. Proyecto que se desarrolla en el marco de los Proyectos De Innovación y Transferencia En Áreas Prioritarias de la Pcia. de Buenos Aires (PIT-AP-BA). Período: 2017-2018.

4. FORMACION DE RECURSOS HUMANOS

En este proyecto se participa en la formación de recursos humanos a través de la dirección de tesis de doctorado, maestría y trabajos finales de especialización y tesinas de grado. Al mismo tiempo participan del proyecto, becarios que están realizando sus tesis de postgrado.

En 2017, se han aprobado 2 trabajos de Maestría y 1 de Especialización en el área de Tecnología y Educación, dirigidos por miembros de este proyecto. Hay 1 tesis doctoral vinculada al proyecto, entregada en 2017 y esperando evaluación. Además se han finalizado 2 tesinas de grado dirigidas por miembros del proyecto en relación a los temas aquí presentados.

Hay 1 trabajo de especialización terminado y esperando su evaluación.

5. BIBLIOGRAFIA

- Archuby, F. (2017). Objeto de Aprendizaje sobre modularización. Presentado en el marco

- del concurso de OA de la Facultad de Informática de la UNLP. En evaluación.
- Archuby, F; Sanz, C. & Pesado, P. (2017) Juego serio como actividad de autoevaluación de los alumnos y su integración con un entorno virtual de enseñanza y aprendizaje. Tesina de Grado aprobada en marzo de 2017. Facultad de Informática. UNLP. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/59652>
 - Artola, V. & Sanz, C. (2017) Learning object for the understanding of the operation merge. 2017 Twelfth Latin American Conference on Learning Technologies (LACLO). Pages: 1 – 4. IEEE Conferences. DOI: 10.1109/LACLO.2017.8120953
 - Assinatto, G., Sanz, C. & Gorga, G. (2017). Propuesta de Tesis de Maestría en Tecnología Informática Aplicada en Educación: La integración de Tecnologías de Información y Comunicación en procesos de enseñanza y aprendizaje de nivel universitario. Aprobada en 2017.
 - Cabero Almenara, J. (2015). Reflexiones educativas sobre las tecnologías de la información y la comunicación (TIC). Revista Tecnología, Ciencia y Educación, [S.l.], jun. 2015. ISSN 2444-2887. Disponible en: <http://www.tecnologia-ciencia-educacion.com/index.php/TCE/article/view/27> Fecha de acceso: 07 mar. 2018
 - Castañeda, L.; Dabbagh, N.; Torres-Kompen, R. *Personal Learning Environments: Research-Based Practices, Frameworks and Challenges*. Journal of New Approaches in Educational Research, [S.l.], v. 6, n. 1, p. 1-2, jan. 2017. ISSN 2254-7339. Available at: <<https://naerjournal.ua.es/article/view/229>>. Accedido el: 12 mar. 2018. doi:<http://dx.doi.org/10.7821/naer.2017.1.229>.
 - Castro, J.; Gorga, G. & De Giusti, L. (2017) Extensión del entorno CMRE mediante la implementación de conceptos relacionados a concurrencia y paralelismo. Su utilización en el ámbito educativo. Tesina de grado aprobada en 2017. Facultad de Informática. UNLP.
 - Dabbagh, N., Fake, H. Percepciones de los estudiantes universitarios sobre los Entornos Personales de Aprendizaje a través de un prisma de herramientas digitales, procesos y espacios. Journal of New Approaches in Educational Research. Vol. 6. No. 1. Enero 2017. pp. 30–38 ISSN: 2254-7339 DOI: 10.7821/naer.2017.1.215
 - Del Río, L.; Bucari, N. & Sanz, C. (2017) Geometría dinámica en entornos hipermedia como facilitadora del aprendizaje de la Matemática. Diseño, implementación y evaluación de un prototipo para el inicio del nivel universitario. Tesis de Maestría de Tecnología aplicada en Educación aprobada en junio de 2017. Facultad de Informática. UNLP
 - Dieser, P; Sanz, C. & Zangara, A. (2017) Estrategias de autorregulación del aprendizaje en escenarios educativos mediados por tecnologías de la información y la comunicación. Una revisión y análisis en la Educación Superior Iberoamericana. Trabajo final de Especialización en TIAE.
 - Digiani, P.; Sanz, C. & Gorga, G. (2017) Propuesta de tesina de grado. *InfoMeeting*: una herramienta de comunicación sincrónica moderada. Aprobada en abril de 2017. Facultad de Informática. UNLP.
 - Estrebou, C.; Salazar, N. & Sanz, C. (2017) Objeto de aprendizaje para la enseñanza de compuertas lógicas: experiencia y evaluación. En Actas del XII Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2017). ISBN: 978-987-44-1704-6. págs. 361-372. Junio de 2017.
 - Fachal A., Sanz C., Abásolo M.J. (2017). Informe técnico sobre “Experiencias desarrolladas en entornos virtuales 3D orientadas a personas con discapacidad auditiva”.
 - Fierro, W.; Sanz, C & Zangara, A. (2017) Estrategias basadas en analítica del aprendizaje para el estudio de las variables relacionadas al aprendizaje autónomo. Estudio de caso en la Educación Superior de Ecuador. Propuesta de Tesis Doctoral aprobada en 2017.
 - Galimberti, S.; Sánchez, B; Gorga, G. & Sanz, C. (2017) Accesibilidad web. Aplicación a un estudio de caso. Tesina de Grado aprobada en 2017. Facultad de Informática. UNLP.
 - Griffin, P., McGaw, B. & Care, E. (Eds.) (2012). *Assessment and teaching of 21st century skills*, Springer, Dordrecht.
 - Herrera, S.; Palavecino, R & Sanz, C. (2017) Aprendizaje de estructuras de datos mediante mlearning. Simposio. I Simpósio Ibero-Americano de Tecnologias Educacionais. ISSN 2594-388X. Págs. 267 – 277. Brasil. Santa Catarina. 2017. Disponible en https://publicacoes.rexlab.ufsc.br/index.php/site_d
 - Iglesias, L. & Archuby, F. (2017) Informe Técnico sobre el Entorno IDEAS.
 - Jaggars, S. & Xu, D. (2016). How do online course design features influence student performance? Computers & Education. Volume 95, 270-284.
 - Kuhn, Caroline. *Are Students Ready to (re)-Design their Personal Learning Environment?*

- The Case of the E-Dynamic.Space. Journal of New Approaches in Educational Research, [S.l.], v. 6, n. 1, p. 11-19, jan. 2017. ISSN 2254-7339. Disponible en: <https://naerjournal.ua.es/article/view/185>. Accedido el 12 mar. 2018 en <http://dx.doi.org/10.7821/naer.2017.1.185>
- Lobos, E.; Gibelli, T. & Sanz, C. (2017) Tecnologías Innovadoras Como Mediadoras de Procesos de Enseñanza y Aprendizaje. Exploración de Herramientas de Realidad Aumentada. Workshop de Investigadores de Ciencia de la Computación. Argentina. Buenos Aires. 2017.
 - Pietroboni, F. & Sanz, C. (2017). Análisis de la implementación de MOOC en educación superior en América Latina. Propuesta aprobada de Trabajo de Especialización en Tecnología Informática Aplicada en Educación.
 - Salazar, N.; Sanz, C. & Gorga, G. (2018). Análisis comparativo de librerías de realidad aumentada. Sus posibilidades para la creación de actividades educativas. Propuesta de Especialización en Tecnología Informática aplicada en Educación aprobada en 2018.
 - Santillan, C.; Sanz, C. & Gorga, G. (2017) ComuniCA: Un componente de software para la escritura de mensajes de Comunicación Aumentativa en espacios web. TEyET 2017. Pp. 340-352. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/63438>
 - Sanz, C. & Barranquero, F. (2017). Material presentado en el curso de Diseño y Producción de Objetos de Aprendizaje, dictado en el marco del Doctorado en Ciencias Informáticas de la UNLP.
 - Sanz, C., Barranquero, F. & Moralejo, L. (2016). CROA: a learning object design and creation methodology to bridge the gap between educators and reusable educational material creation. EDULEARN 2016 - 8th annual International Conference on Education and New Learning Technologies. Barcelona (Spain). 4th - 6th of July, 2016. Pages: 4583-4592. ISBN:978-84-608-8860-4 ISSN:2340-1117. [10.21125/edulearn.2016.2101](https://doi.org/10.21125/edulearn.2016.2101)
 - Sanz, C., Zangara, A. & Dieser, P. (2016). Modelo de seguimiento de una actividad de foro en la enseñanza universitaria. Actas del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016), ISBN: 978-987-733-072-4, págs. 431-440.
 - Sanz, C.; Barranquero, F. & Astudillo, G. (2017) El concepto de OA bajo una mirada techno-educativa. Debates y tendencias. Tutorial dictado en el marco de LaClo 2017. La Plata. 2017.
 - Sanz, C; Zangara, A. & Dieser, P. (2016) Modelo de seguimiento de una actividad de foro en la enseñanza universitaria. Aplicación a un estudio de caso. XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016). San Luis. Octubre de 2016. Páginas 431-440.
 - Simón, L; Zangara, A. & Gonzalez A. (2017) La enseñanza de una lengua extranjera utilizando tecnología digital: estudio descriptivo del tránsito de un material impreso a un entorno virtual de aprendizaje. Tesis de Maestría en Tecnología Informática aplicada en Educación aprobada en junio de 2017. <http://sedici.unlp.edu.ar/handle/10915/61340>
 - Violini, L. & Sanz, C. (2017) Learning objects: How to insert an element into a vector? Pp: 1-4. DOI: 10.1109/LACLO.2017.8120956
 - Violini, L.; Sanz, C. & Pesado, P. (2017). Propuesta de un *framework* para la creación de objetos de aprendizaje. Actas del XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017). ISBN: 978-950-34-1539-9, págs. 383-393, octubre de 2017.
 - Zangara, A. & Sanz, C. (2017) "Displaying the collaborative process as meta-knowledge. Description of a mirroring strategy and its results". En: Communications in Computer and Information Science (CCIS) vol. 790. P. 79-89. Springer.

Tendencias en interacción persona-ordenador. Aplicaciones en escenarios educativos

Sanz Cecilia^{1,3}, Gorga Gladys¹, Artola Verónica^{1,5}, Moralejo Lucrecia¹, Salazar Mesía Natalí^{1,6}, Archuby Federico¹, Nordio Mauricio¹, Abásolo María José³, Baldassarri Sandra², Manresa-Yee Cristina⁴, Pesado Patricia^{1,3}

¹ Instituto de Investigación en Informática LIDI (III-LIDI). Centro Asociado CIC
Facultad de Informática – Universidad Nacional de La Plata

² GIGA AffectiveLab, Universidad of Zaragoza, España

³ Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

⁴ Departamento de Ciencias Matemáticas e Informática. Universidad de las Islas Baleares, España

⁵ Becaria doctoral CONICET

⁶ Becaria TIPO A UNLP

{[csanz](mailto:csanz@unlp.edu.ar), [ppesado](mailto:ppesado@unlp.edu.ar), [ggorga](mailto:ggorga@unlp.edu.ar), [vmartola](mailto:vmartola@unlp.edu.ar), [lmoralejo](mailto:lmoralejo@unlp.edu.ar), [nsalazar](mailto:nsalazar@unlp.edu.ar), [farchuby](mailto:farchuby@unlp.edu.ar)}@lidi.info.unlp.edu.ar, sandra@unizar.es, cristina.manresa@uib.es

RESUMEN

En este trabajo se presenta una línea de investigación, desarrollo e innovación en modelos y paradigmas de interacción persona ordenador y sus posibilidades para el escenario educativo. Se abordan estudios teóricos, revisión sistemática de la literatura, se desarrollan aplicaciones utilizando realidad aumentada, interacción tangible e hibridaciones de estos paradigmas de interacción. Al mismo tiempo, se analiza los entramados entre mundo digital y físico, de manera tal de favorecer procesos educativos. De manera especial, se incluyen en esta línea los juegos serios utilizando diferentes modelos de interacción.

Se presentan en este trabajo los principales resultados alcanzados en relación a esta línea temática en el año 2017.

Palabras claves: Realidad Aumentada, Interacción Tangible, Computación Afectiva, Dispositivos móviles.

CONTEXTO

Esta investigación forma parte del subproyecto “Escenarios educativos mediados por Tecnologías de la Información y la Comunicación (TIC)”, acreditado por el Ministerio de Educación de la Nación. Este proyecto continúa en 2018 con un nuevo subproyecto denominado “Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios

educativos híbridos”, enmarcado en el proyecto “Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso”.

1. INTRODUCCION

La evolución de las interfaces y de los modelos de interacción ha permitido acercar a las diferentes disciplinas, y a las personas, en general, al uso cotidiano de herramientas digitales. Las interfaces de usuario naturales (NUI: *Natural User Interfaces*) y la computación ubicua han empezado a trazar nuevos caminos en esta relación entre las personas, el mundo físico y el digital (Duque & Vásquez, 2015).

La posibilidad de combinar el uso de materiales de aprendizaje físicos con las posibilidades del entorno digital, han generado gran expectativa entre los académicos y también en el escenario educativo. La larga historia en el mundo de la educación en relación a la utilización de objetos físicos y manipulables para el aprendizaje desde las etapas iniciales de la vida de un niño ya forman parte de nuestra cultura (Cuendet, Bumbacher & Dillenbourg, 2012). En las dos últimas décadas se ha visto un incremento en la exploración del uso de tecnologías digitales para la enseñanza y el aprendizaje, con un reciente énfasis en la interacción multisensorial y multimodal (Price, Duffy & Gori, 2017). Sin embargo, es necesario reflexionar acerca de: ¿qué ventajas ofrecen estos escenarios híbridos? ¿Qué objetos del mundo físico y cuáles del digital integrarán la experiencia

educativa? ¿Qué relación guardarán entre ellos? ¿De qué manera los estudiantes y docentes participarán e interactuarán con dichos objetos? ¿Qué competencias pueden desarrollarse a través de este tipo de actividades y experiencias? ¿Cuándo es necesario apelar a la multimodalidad y a la multisensorialidad?

En esta línea de investigación se abordan este tipo de interrogantes. Se realizan estudios teóricos y experimentales sobre las posibilidades de diferentes tecnologías emergentes y paradigmas de interacción como la realidad aumentada, la interacción tangible, y la computación ubicua en general. Se diseñan experiencias en mundos virtuales 3D, se estudian los fundamentos de la computación afectiva que podrían enriquecer los escenarios educativos híbridos y se abordan y analizan juegos serios con diferentes tecnologías, y más específicamente con dispositivos móviles.

2. TEMAS DE INVESTIGACION / DESARROLLO ABORDADOS

Las líneas de investigación, desarrollo e innovación en las que se trabaja son:

- Estilos y Paradigmas de Interacción Persona – Ordenador. Posibilidades para la educación, y con foco en la accesibilidad e inclusión.
- Realidad Virtual y Aumentada en escenarios educativos.
- Interacción tangible y multitáctil. Aplicaciones en educación y educación especial.
- Computación afectiva para el enriquecimiento de escenarios educativos.
- Dispositivos móviles y sensores para la generación de entornos aumentados.
- Juegos serios con diferentes dispositivos y modelos de interacción.

3. RESULTADOS ESPERADOS/OBTENIDOS

En el año 2017 se ha trabajado en las diferentes temáticas que se abordan en esta línea de investigación. El equipo se compone con investigadores formados, en formación y becarios, al mismo tiempo se nutre del aporte y la cooperación con investigadores formados de otras universidades,

Se describen aquí diferentes acciones realizadas en este período.

En 2017 se realizó un estudio bibliográfico sobre experiencias de realidad aumentada en espacios educativos, además se comenzaron a analizar diferentes librerías y herramientas para la creación de contenidos de realidad aumentada (RA) (Salazar, Sanz & Gorga, 2018). Se diseñó una guía para el diseño de actividades educativas basadas en RA (Sanz, 2017). Al mismo tiempo, se desarrollaron experiencias con docentes utilizando las herramientas estudiadas. Se recopilaron también aplicaciones móviles que utilizan RA y se las presentó en un taller dictado en la Universidad Nacional de Río Negro (Sanz & Gibelli, 2015), actividad que se desarrolló como parte de un proyecto de investigación conjunto. Se avanzó en el marco del proyecto Reforticca en la propuesta de una aplicación para móviles para ofrecer información sobre los componentes nutricionales de los alimentos en venta en los supermercados y comercios de la región, y también sobre las posibilidades de reciclaje de sus envases, de manera de aportar al consumo responsable (Boza, Prinsich, Abásolo & Sanz, 2017). Además se prevé profundizar en la cooperación en estos temas con la Universidad de las Islas Baleares, en 2018 se realizará una estancia de investigación en el marco de un postdoc para profundizar en temas de RA.

Por otra parte, se continúa trabajando en la investigación en temas relacionados con interacción tangible. Se trabajó en un proyecto con alumnos de grado, en el que se desarrolló FraccionAR. Se trata de un juego serio basado en interacción tangible (ver Figura 1), que permite ejercitar el concepto de fracciones, y sus diferentes representaciones. El juego propone una dinámica competitiva entre dos participantes o dos equipos, aunque puede ser jugado de manera individual (Nordio, Artola & Sanz, 2017). En este caso se integró el uso de objetos físicos (porciones de pizza de juguete y de chocolates) para poder interactuar con la aplicación informática sobre una mesa aumentada computacionalmente (*tabletop*). De esta forma, se buscó facilitar en alumnos de nivel inicial un concepto abstracto como el de fracciones. Durante 2018 se llevarán adelante experiencias en escuela utilizando FraccionAr.



Figura 1. Juego FraccionAR presentado en la Expo Universidad y en la exposición de Ciencia y Tecnología de la F. de Informática, UNLP.

Se publicó un trabajo sobre el proceso de evaluación de ACoTI, basado en una revisión bibliográfica en profundidad (Sanz et al., 2017). Este trabajo ha sido fruto de un proceso de varios años de investigación. ACoTI es un desarrollo conjunto con la Universidad de Zaragoza, se trata de una aplicación conformada por una serie de juegos de asociación que se orientan a personas con necesidades complejas de comunicación (Sanz et al., 2013). Este proceso de evaluación motivó el desarrollo de TIES, un editor para la creación de actividades de interacción tangible (Artola, Sanz, Moralejo, Pesado & Baldassarri, 2015), durante 2017 se realizaron las primeras evaluaciones de este editor y serán publicadas durante 2018. Se continuó con el desarrollo de experiencias utilizando ITCol (un juego basado en interacción tangible para la colaboración), que se integra como parte de un seminario de la Maestría en Tecnología Informática Aplicada en Educación de la Facultad de Informática de la UNLP. Para evaluar las sesiones con ITCol se continúa

trabajando con la metodología diseñada por los autores de este trabajo y presentada en (Zangara & Sanz, 2015).

En 2017 se desarrollaron nuevas experiencias utilizando Desafiarte un juego de preguntas y respuestas vinculado a un entorno virtual de enseñanza y aprendizaje (EVEA) IDEAS¹. Este juego permite que los docentes publiquen las autoevaluaciones generadas en el EVEA como una aventura dentro del juego, que presenta desafíos a los participantes. Se llevaron adelante experiencias en un curso del primer año de las carreras de la Facultad de Informática de la UNLP para la autoevaluación de los alumnos previo a rendir examen parcial. Se analizó la motivación, el mejoramiento del estrés que implica una autoevaluación a través del juego y aspectos de mejora propuestos (Archuby, Sanz & Pesado, 2017). Se continuará con este trabajo en 2018. También en esta línea se ha desarrollado un juego serio para la enseñanza de la Programación, llamado AstroCódigo² (Bione, Miceli, Sanz & Artola, 2017). Se está trabajando con dicho juego en los cursos de ingreso a la Facultad de Informática, y se han elaborado instrumentos específicos para indagar sobre el impacto en el aprendizaje y la motivación. En relación a la temática de mundos virtuales 3D, se han abordado diferentes tesis. Se ha avanzado con el estudio de las posibilidades de estos mundos para personas con discapacidad auditiva (Fachal, Sanz & Abásolo, 2017) y se han desarrollado una serie de escenarios en OpenSim para llevar adelante el trabajo experimental de la tesis durante el 2018.

En relación a los temas vinculados con computación afectiva, se han presentado dos propuestas de doctorado que han sido aprobadas. Una relacionada con realizar análisis de sentimientos a partir de mensajes escritos en espacios de comunicación de cursos en línea (Rodríguez, Baldassarri & Sanz, 2017). La segunda propuesta se relaciona con la integración de un componente que permita la recuperación y detección de emociones en entornos virtuales de enseñanza y aprendizaje. Ambas dirigidas por miembros del proyecto en conjunto con la Universidad de Zaragoza.

¹ IDEAS: <http://ideas.info.unlp.edu.ar> – Es una evolución del EVEA WebUNLP, desarrollado en la Facultad de Informática de la UNLP.

² AstroCódigo: <http://www.astrocodigo.com.ar>

En cuanto a los proyectos y acuerdos de cooperación vinculados con la temática, el III- LIDI participa en los siguientes:

- Acuerdo de colaboración en estos temas con la Universidad de Zaragoza. En particular con el grupo GIGA *AffectiveLab*.
- Acuerdo de colaboración en estos temas con la Universidad de las Islas Baleares. En particular con la Unidad de Gráficos, Visión por Computador e IA.
- REDAUTI: Red Temática en Aplicaciones y Usabilidad de la Televisión Digital Interactiva.
- REFORTICCA: Recursos para el Empoderamiento de FORMadores en TIC, Ciencias y Ambiente. Proyecto que se desarrolla en el marco de los Proyectos De Innovación y Transferencia En Áreas Prioritarias de la Pcia. de Buenos Aires (PIT-AP-BA). Período: 2017-2018.
- Acuerdo de cooperación con la Universidad Nacional de Santiago del Estero. Se asesora en temas relacionados a la integración de dispositivos móviles en procesos educativos.
- Proyecto de colaboración con la Universidad Nacional de Río Negro.

4. FORMACION DE RECURSOS HUMANOS

En el marco del subproyecto que aquí se presenta se participa en la formación de los siguientes recursos humanos:

- Dos becarios de UNLP con beca TIPO A, que está realizando la tesis de maestría en TIAE relacionada con la temática aquí expuesta, y trabajan en el marco del subproyecto.
- Un becario doctoral CONICET, que está realizando su tesis en temas vinculados con Interacción Tangible para el escenario educativo, también es integrante del subproyecto.
- Dos especialistas en TIAE, que están finalizando su tesis de maestría y son integrantes del subproyecto.

Los investigadores formados de este proyecto se encuentran dirigiendo tesis doctorales y de maestría vinculadas a los temas de esta investigación, así como también tesinas de grado.

En la siguiente sección se presenta bibliografía de referencia para la investigación en los temas presentados y los trabajos citados a lo largo del documento.

5. BIBLIOGRAFIA

- Archuby, F. H., Sanz, C. V., Pesado, P. M. (2017). Desafiate: juego serio para la autoevaluación de los alumnos y su integración con un entorno virtual de enseñanza y aprendizaje. Actas del XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017). ISBN: 978-950-34-1539-9, págs. 284-294, octubre de 2017.
- Artola, V., Sanz, C. V., Moralejo, L., Pesado, P. M. & Baldassarri S. (2015). Herramienta de autor para la creación de actividades educativas basadas en Interacción Tangible. En XXI Congreso Argentino de Ciencias de la Computación. Junín, 2015.
- Bione, J., Miceli, P., Sanz, C., Artola, V. (2017). Astrocode: A Serious Game For The Development Of Computational Thinking Skills. 9th International Conference on Education and New Learning Technologies.: IATED. vol. n°. p4624 - 4634. issn 2340-1117. 2017
- Bione, J., Miceli, P.; Sanz, C., Artola, V. (2017). Enseñanza de la programación con astrocódigo. Actas del XII Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2017), ISBN: 978-987-44-1704-6, págs. 454-455, junio de 2017.
- Boza, R., Prinsich, N., Abásolo M. J., Sanz, C. (2017). Aplicación móvil para el consumo consciente y responsable de productos alimenticios. Propuesta de Tesina de grado de Licenciatura de la Facultad de Informática de la UNLP aprobada en 2017.
- Cuendet, S., Bumbacher, E., Dillenbourg, P. (2012). Tangible vs. virtual representations: When tangibles benefit the training of spatial skills. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (pp. 99–108). ACM.
- Duque, E., Vásquez, A. (2015). NUI para la educación. Eliminando la discriminación

- tecnológica en la búsqueda de la Inclusión digital. Disponible en: <http://hdl.handle.net/123456789/3780>
- Fachal, A., Sanz, C., Abásolo, M.J. (2017). Informe técnico sobre “Experiencias desarrolladas en entornos virtuales 3D orientadas a personas con discapacidad auditiva”.
 - Nordio, M., Artola, V., Sanz, C. (2017). FraccionAR: juego sobre fracciones basado en Interacción Tangible presentado en el marco de la convocatoria de proyectos de innovación de la Facultad de Informatics 2017. Muestra del Trabajo realizado en la ExpoCiencia 2017 y la ExpoUNLP disponible en: <http://weblidi.info.unlp.edu.ar/catedras/proyectoAlu/video.mp4>
 - Price, S., Duffy S., Gori, M. (2017). Developing a Pedagogical Framework for Designing a Multisensory Serious Gaming Environment. Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education. ISBN: 978-1-4503-5557-5. <http://doi.acm.org/10.1145/3139513.3139517>
 - Rodriguez, V., Baldassarri, S., Sanz, C. (2017). Informe de avance sobre la tesis Reconocimiento de información afectiva en EVEAS a partir de espacios de interacción textual. Tesis doctoral con propuesta aprobada.
 - Salazar Mesía, N., Sanz, C., Gorga, G. (2018) Análisis comparativo de librerías de Realidad Aumentada. sus posibilidades para la creación de actividades educativas. Propuesta de Especialización de Tecnología Informática Aplicada en Educación aprobada en marzo 2018.
 - Sanz, C., Guisen, A., De Giusti, A., Baldassarri, S., Marco, J.; Cerezo, E. (2013). Games as educational strategy: A case of tangible interaction for users of Alternative and Augmentative Communication International Conference on Collaboration Technologies and Systems (CTS), San Diego, CA, 2013, pp. 377-381. doi: 10.1109/CTS.2013.6567258
 - Sanz, C., Gibelli, T. (2015). La mediación de las tecnologías de la información y la comunicación en procesos educativos. Innovaciones para mejorar los procesos de enseñanza y aprendizaje. Aprobado en el marco de la Convocatoria BIENAL PI-2015. Universidad Nacional de Rio Negro.
 - Sanz, C. (2017). Guía de diseño de actividades educativas basadas en Realidad Aumentada. Desarrolla en el marco del proyecto F-11/016 2017.
 - Sanz, C., Artola, V., Guisen, A.; Marco, J.; Cerezo, E.; Baldassarri, S. (2017) Shortages and Challenges in Augmentative Communication through Tangible Interaction Using a User-centered Design and Assessment Process. Journal of Universal Computer, volume 23, issue 10. Disponible en: http://www.jucs.org/jucs_23_10/shortages_and_challenges_in
 - Zangara, A., Sanz, C. (2015). Collaborative Learning Work. Evaluation Model Using The Ethnographic Methodology Of “Analysis Of Social Networks”. 7th International Conference on Education and New Learning Technologies .EDULEARN 2015. España-Barcelona. Pp. 5373-5380. ISBN: 978-84-606-8243-1 - ISSN: 2340-1117 – Editor: IATED. Disponible en: <http://library.iated.org/view/ZANGARA2015COL>

Posibilidades de los dispositivos móviles para la educación en colegios de nivel secundario

Pizarro Rubén, Testa Oscar, Camiletti Pablo, Ascheri María Eva, Lobos Martín, Perez Ramiro, Rojas Alejandro

Departamento de Matemática / Facultad de Ciencias Exactas y Naturales/ Universidad Nacional de La Pampa

Uruguay 151, 00-54-02954-245220 int 7124
mavacheri@gmail.com ; otesta@gmail.com

RESUMEN

Los dispositivos móviles y el acceso a las redes modifican el entorno de estudiantes, profesores y personal del sistema educativo, sabemos cómo han provocado profundos cambios en “La Sociedad Red” y han afectado el mencionado sistemas educativos.

En un primer momento del proyecto se espera realizar una revisión bibliográfica que aborde el tema de investigación, las aplicaciones existentes destinadas a la enseñanza en el nivel de la educación secundaria y las experiencias de inclusión de dispositivos móviles en este nivel. Posteriormente definiremos las instituciones, el diseño y elaboración de herramientas para la recopilación de información para avanzar con la experiencia directa en el ámbito educativo y la recopilación de la información propiamente dicha. Finalmente se espera organizar, analizar y elaborar conclusiones.

En este proyecto nos proponemos analizar la compleja influencia del acceso a las redes y la utilización de dispositivos móviles y cómo aprovechar estas características para experimentar, analizar y proponer mejoras en las formas de utilizar estas tecnologías en los colegios de nivel secundario de la Ciudad de Santa Rosa

Palabras clave: dispositivos móviles, aplicaciones, enseñanza aprendizaje.

CONTEXTO

El proyecto de investigación que se presenta está radicado en el Departamento de Matemática de la Facultad de Ciencias Exactas y Naturales de la UNLPam y financiado por dicha Facultad.

Tiene una duración de tres años, siendo este el segundo año de su desarrollo. Está vinculado con otros proyectos del Departamento ya ejecutados y en ejecución, relacionados con la enseñanza en nivel secundario y la inclusión de tecnologías en educación.

1. INTRODUCCIÓN

Como leemos en la mayoría de los trabajos y compartiendo sin dudas esa afirmación, podemos decir que la presencia de los dispositivos móviles y su vertiginoso avance han modificado muchas de las actividades que se desarrollan en nuestra sociedad. Estos dispositivos que surgieron como herramientas de comunicación y entretenimiento han llegado a desarrollar un papel fundamental en el mundo de la economía y el conjunto de la sociedad (Shuler et al, 2013).

Sin dudas los dispositivos móviles están presentes en la vida de la mayoría de docentes y estudiantes para realizar diversas actividades. Ante esta realidad los Ministerios de Educación en diferentes países han iniciado políticas tendientes a que estos dispositivos cobren mayor importancia también en educación. En esta línea podemos mencionar los modelos 1 a 1 implementados en muchos países que consisten en la provisión por parte del estado de un dispositivo móvil (netbook o tableta) a cada alumno y docente. También existen iniciativas del tipo “Trae tu propia tecnología” denominada (BYOT, por sus siglas en inglés ‘Bring Your Own Technology’), consistentes en que cada estudiante lleve su propio dispositivo móvil para utilizarlo en la escuela. Esta metodología se ha implementado en sectores sociales más desarrollados en los cuales todos pueden acceder a estos dispositivos.

Los proyectos mencionados anteriormente, entre otros de similares características, permiten sin dudas un mayor acceso a la tecnología y aumenta las posibilidades de acceso a la información de los docentes y estudiantes. Permitirían así acceder a las principales características tecnológicas del aprendizaje móvil que según indica Cantillo Valero, et. al. (2012), son las siguientes:

- Portabilidad, debido al pequeño tamaño de los dispositivos.

- Inmediatez y conectividad mediante redes inalámbricas.
- Ubicuidad, ya que se libera el aprendizaje de barreras espaciales o temporales.
- Adaptabilidad de servicios, aplicaciones e interfaces a las necesidades del usuario. También existe la posibilidad de incluir accesorios como teclados o lápices para facilitar su uso.

Vivimos indudablemente momentos en los cuales las tecnologías avanzan vertiginosamente pero existe una amplia brecha entre el desarrollo y la implementación (Montoya, 2009); es así que continuando con el análisis de este autor, nos indica que para lograr la inclusión de los dispositivos móviles, en especial los teléfonos inteligentes, se puede temporalizar tres pasos:

1. Apropiación del objeto
2. Apropiación de la funcionalidad
3. Apropiación de la nueva forma de aprendizaje

Las dos primeras etapas como señala Díaz Bilbao (2014), serían alcanzadas rápidamente por los estudiantes, quedando la tercera etapa a cargo de las instituciones educativas y los docentes.

Sin duda la mayoría de los autores coinciden en indicar que la inclusión de dispositivos móviles en el aula, produce grandes beneficios: motivan a los estudiantes, aumenta las habilidades sociales mejorando la cooperación y la colaboración mejorando la creatividad y la capacidad cognitiva.

Sin embargo como ha sucedido con otras tecnologías, percibimos que los dispositivos móviles actualmente no son tenidos en cuenta ni aprovechados en las formas ni en la magnitud que observamos sería conveniente. Por el contrario en algunos ámbitos han sido prohibidos por causar distracción y otros problemas entre los estudiantes y la comunidad educativa.

Como lo expresan Castells, Palamidessi y muchos otros, los cambios que se han

producido en las formas de comunicarse, acceder a información y conocer el mundo, mucho tienen que ver el acceso a las redes y los nuevos dispositivos y esto pareciera ser de alguna manera todavía resistido en el ámbito educativo. Al mismo tiempo se reconoce una crisis identificada en las dificultades que se presentan con la falta de atención, falta de aprendizaje o productividad, de motivación, de malestar de estudiantes y profesores entre otros problemas.

¿Tendremos que plantearnos abordar el trabajo en las instituciones educativas desde nuevos paradigmas? Pensar alternativas para las instituciones educativas implica salir de lugares comunes, observar que el mundo y la sociedad han cambiado mucho, en comparación las instituciones educativas, emblemáticas a la hora de hablar del conocimiento, parecen no haber dado signos claros de entender la metamorfosis del conocimiento a la que hoy asistimos.

Es evidente que este contexto no puede ser abordado con propuestas simplistas, pensar que la sola inclusión de dispositivos móviles puede generar las condiciones necesarias para mejorar las posibilidades de enseñanza y aprendizaje en la educación no alcanza.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación que se abordan con el presente proyecto son las siguientes:

1. Nivel de penetración de los dispositivos móviles en las actividades de enseñanza aprendizaje en aulas de nivel medio.
2. Búsqueda, experimentación y análisis de aplicaciones educativas para la enseñanza en el nivel medio
3. Características metodológicas de las clases de nivel medio en las que se incluye dispositivos móviles.

3. RESULTADOS OBTENIDOS / ESPERADOS

Los objetivos planteados son los siguientes:

- Realizar una revisión bibliográfica de experiencias de inclusión y características de los dispositivos móviles en el ámbito educativo.
- Indagar formas y frecuencia de la utilización por parte de profesores de los dispositivos móviles en el nivel secundario de Santa Rosa.
- Indagar expectativas de los profesores ante la aparición de los dispositivos móviles en el nivel secundario de Santa Rosa.
- Analizar cómo se utilizan los dispositivos móviles en las actividades de enseñanza.
- Analizar cómo la utilización de dispositivos móviles en las actividades de enseñanza modifica las formas de evaluar a los estudiantes.
- Analizar el nivel de formación del profesorado en el uso de dispositivos móviles.
- Búsqueda y recopilación de aplicaciones educativas.
- Experimentar, en distintos centros educativos de nivel secundario, actividades utilizando aplicaciones para dispositivos móviles.

Se espera que a partir de las actividades de investigación desarrolladas en este proyecto, se pueda en primer lugar, tener un análisis de la situación actual a nivel de dispositivos disponibles en los colegios de la ciudad de Santa Rosa, normativas legales, infraestructura tecnológica, conocimientos y habilidades de los docentes y estudiantes en el uso de estas tecnologías.

Con estas experiencias se espera encontrar características que los dispositivos móviles pueden aportar para la mejora del entorno educativo, en particular la enseñanza y el

aprendizaje, siempre utilizando los equipos que poseen docentes y estudiantes.

También se espera desarrollar acciones de capacitación de docentes de nivel medio sobre la utilización de estas tecnologías en las aulas, de manera de extender y profundizar las mejoras que los dispositivos móviles y el tipo de actividades asociados a ellos puedan aportar al ámbito educativo.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo es de carácter interdisciplinario, está conformado con especialistas del área de Educación, Matemática y Computación. Participan además dos estudiantes del Profesorado en Computación.

Se realizan, también, actividades de capacitación y transferencia referidas a la utilización de aplicaciones educativas para dispositivos móviles con sistema operativo Android.

5. BIBLIOGRAFÍA

- Argentina, Ministerio de Educación de Argentina. (2015). *Conectar Igualdad, fundamentos del programa*. Disponible en <http://www.conectarigualdad.gob.ar/seccion/sobre-programa/fundamentos-del-programa-17>
- Castells, M. (2016). *¿Comunidades virtuales o sociedad en red?*. Plaza & Janés Editores. Barcelona, España
- Castells, M., & Schmalenberger, M. (2006). *La sociedad red: una visión global* (No. U10 1087). IICA, Bs As (Argentina).

- Cantillo Valero, C., Roura Redondo, M. & Sánchez Palacín, A. (2012). Tendencias actuales en el uso de dispositivos móviles en educación. Educational Portal of the Americas – Department of Human Development, Education and Culture. OEA-OAS ISSN 0013-1059 La Educación Digital Magazine N 147 – www.educoas.org
- Díaz Bilbao (2014) El smartphone como herramienta educativa. Trabajo final de maestría. Máster en formación del profesorado de educación secundaria (ámbito tecnológico). Universidad Pública de Navarra.
- Montoya, M. (2009). Mobile learning – mlearning- technology resources and their relationship with distance learning environments: applications and research studies. Tecnológico de Monterrey, ITESM (México).
- Miranda Ruiz, E. (2010). *Paradigma Interpretativo en Investigación*. Disponible <http://www.monografias.com/trabajos97/paradigma-interpretativo/paradigma-interpretativo.shtml>
- Palamidessi, M. (2006). La escuela en la sociedad de redes. *Una introducción a las tecnologías de la información y la comunicación*.
- RINALDI, M. (2012). *Revolución Mobile Learning*. Disponible en: <http://es.slideshare.net/crossmedialab/revolucion-mobile-learning> Consultado 06/08/2013.
- Shuler C, Niall Winters & Mark West. (2013) El futuro del aprendizaje móvil. Implicaciones para la formulación y planificación de políticas et al. UNESCO working paper series on mobile learning. Disponible en <http://unesdoc.unesco.org/images/0021/002196/219637s.pdf> Consultado 10/08/16.

Vinculación de líneas de I+D con la extensión universitaria. El caso de GEMIS en UTN-FRBA

Straccia, Luciano; Pytel, Pablo; Pollo Cattaneo, Ma. Florencia

Grupo de Estudio en Metodologías de Ingeniería de Software (GEMIS)
Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.

Medrano 951 (C1179AAQ) Ciudad Autónoma de Argentina. Buenos Aires Tel +54 11 4867-7511

lstraccia@frba.utn.edu.ar, flo.pollo@gmail.com, ppytel@gmail.com

RESUMEN

En este trabajo se presenta, con una mirada integral, las actividades del Grupo de Estudio de Metodologías para Ingeniería en Software y Sistemas de Información (GEMIS) de la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional, se enuncian y describen sus líneas de trabajo, sus proyectos y las actividades de vinculación entre los mismos.

Los diferentes proyectos de investigación y desarrollo y, las actividades relacionadas a otros grupos de trabajo de la institución en la cual lleva adelante sus actividades, configuran una cartera de proyectos que deben ser gestionados integralmente. Esto requiere la identificación de componentes y actividades vinculables entre cada uno de los proyectos y la gestión eficiente de su puesta en funcionamiento y control.

Palabras clave: portafolio de proyectos, proyectos I+D, extensión universitaria, sistemas inteligentes, tecnología educativa, gestión del conocimiento.

CONTEXTO

En marzo de 2009 y en el ámbito de la Facultad Regional Buenos Aires de la

Universidad Tecnológica Nacional (UTN-FRBA), se conformó el Grupo de Estudio de Metodologías para Ingeniería en Software y Sistemas de Información (GEMIS). Desde ese momento y sostenidamente, se integró un equipo de docentes y alumnos con interés en la investigación, la sistematización de cuerpos de conocimiento y la promoción sobre el campo de la Ingeniería en Sistemas de Información y la Ingeniería en Software, sus aplicaciones y abordajes metodológicos en todo tipo de escenarios convencionales y no convencionales.

1. INTRODUCCIÓN

Una cartera o portafolio de proyectos es "una colección de componentes (proyectos, programas, y otro trabajo tal como mantenimiento y operaciones en curso) que se agrupan para facilitar la gestión efectiva de ese trabajo, en busca de alcanzar los objetivos estratégicos. Los proyectos o programas del portafolio pueden no ser necesariamente independientes o directamente relacionados" [1].

La gestión de portafolio de proyectos (habitualmente conocida como PPM, sus siglas en inglés provenientes de los términos Project Portfolio Management) es "la administración centralizada de una

o más carteras, e implica identificar, priorizar, autorizar, administrar y controlar proyectos, programas y otros trabajos relacionados" [2].

GEMIS organiza sus actividades de investigación, desarrollo e innovación a través de diferentes proyectos vinculados a las siguientes líneas de trabajo: Ingeniería de Software / Ingeniería de Requisitos; Educación y Tecnología; Sistemas Inteligentes / Inteligencia Artificial; Explotación de Información / Base de Datos; Arquitectura e Infraestructura; Gestión del Conocimiento.

GEMIS articula sus actividades con las cátedras de Sistemas y Organizaciones e Inteligencia Artificial, asignaturas pertenecientes a la carrera de Ingeniería en Sistemas de Información de la UTN-FRBA, con la Especialización y la Maestría en Ingeniería en Sistemas de Información de la Escuela de Posgrado de UTN-FRBA y con proyectos de la Secretaría de Cultura y Extensión Universitaria (SCEU) de la UTN-FRBA. De esta forma, GEMIS busca también realizar aportes sustantivos a los objetivos de la Universidad: enseñanza (formación de profesionales y científicos), investigación (producción de conocimiento científico) y extensión (producción de servicios de asistencia técnica, consultoría y transferencia tecnológica al sector productivo y a la comunidad) [3].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta sección se presentan los diferentes proyectos de investigación y desarrollo y los proyectos de extensión vinculados que forman parte de la actividad de GEMIS.

Proyecto de investigación y desarrollo (PID): La gestión del conocimiento en pequeñas y medianas fábricas de software en el AMBA

Este proyecto de investigación y desarrollo se inició en Enero de 2018 y tiene vigencia hasta Diciembre 2019. Busca indagar cuál es el grado de implementación de gestión del conocimiento en las pequeñas y medianas fábricas de software del Área Metropolitana de Buenos Aires (AMBA), cuáles son los problemas presentes en esta implementación y qué modelos de gestión del conocimiento se podrían proponer para que las pequeñas y medianas fábricas de software del AMBA pudieran implementarlos.

Proyecto de investigación y desarrollo (PID): Implementación de sistemas inteligentes para la asistencia a alumnos y docentes de la carrera de Ingeniería en Sistemas de Información.

Este proyecto de investigación y desarrollo, iniciado en 2016 y con vigencia hasta fin de 2019, busca implementar un Sistema Inteligente que asista el proceso de enseñanza-aprendizaje entre los alumnos y sus docentes de la asignatura Sistemas y Organizaciones de la UTN-FRBA. Entre sus principales funcionalidades se destacan la evaluación de los alumnos sobre los temas teóricos, la revisión automática de ejercicios prácticos, la generación de recomendaciones a los alumnos sobre los materiales de estudio y la generación de reportes para los docentes sobre el nivel de conocimiento de los alumnos.

Proyecto de investigación y desarrollo (PID): Intervenciones tecnológicas en dispositivos didácticos con herramientas de tecnología informática

Este proyecto de investigación y desarrollo tiene como objetivo describir y analizar el uso de la tecnología informática en las intervenciones didácticas de los profesores de las asignaturas de la carrera de Ingeniería en Sistemas de Información de la UTN-FRBA y desarrollar nuevos artefactos tecnológicos que favorezcan la mejora en las intervenciones didácticas y una metodología de implementación. Se inició en 2016 y si bien su finalización formal fue en Diciembre 2017, se encuentra transitando la última etapa de implementaciones y evaluación de los resultados.

Proyecto de extensión universitaria: Atender, software para el seguimiento integral de los estudiantes

El objetivo de este proyecto de extensión universitaria, aprobado en el año 2017 por la Secretaría de Políticas Universitarias del Ministerio de Educación en la Convocatoria de Desarrollos para la Innovación Social, con vigencia hasta Agosto de 2018, y desarrollado bajo las actividades de GEMIS, es brindar una herramienta de tecnología informática a los actores del sistema educativo de nivel superior que permita realizar un seguimiento eficiente de los estudiantes. Los docentes contarán con información centralizada de diversas variables y orígenes: asistencia de los alumnos a clases, desempeño en evaluaciones formales y no formales, resultados de encuestas. El software dispondrá de las facilidades para cargar esta información o para ser obtenida

desde otras herramientas que el docente pudiera utilizar y brindará información resumida que permita al docente identificar riesgos de deserción o problemáticas en el desempeño de los alumnos, permitiendo al propio docente tomar decisiones sobre su propia actuación a partir de información objetiva.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el marco del proyecto “Implementación de sistemas inteligentes para la asistencia a alumnos y docentes de la carrera de Ingeniería en Sistemas de Información”, a través de la aplicación de modelos y métodos propios de los sistemas inteligentes se llevaron a cabo trabajos sobre las asignaturas Sistemas y Organizaciones e Inteligencia Artificial cuyos resultados fueron utilizados para la mejora de las asignaturas correspondientes. La presentación de estos resultados fue realizada en [4-6]. En los dos últimos trabajos se presenta el uso de técnicas y herramientas propias de las disciplinas de Explotación de Información e Inteligencia de Negocios. En la última etapa del proyecto se prevé la implementación de un módulo de sistemas inteligentes y su integración con el Proyecto Atender.

Entre las actividades del proyecto “Intervenciones tecnológicas en dispositivos didácticos con herramientas de tecnología informática” se han analizado diversas problemáticas [7] y se han propuesto algunas soluciones tecnológicas [8,9]. Además a partir de algunas dificultades presentadas en el avance del desarrollo del software se ha producido el descubrimiento de nuevas estrategias y metodologías, para el desarrollo de software, vinculadas a las características de los proyectos de

investigación y desarrollo en las instituciones de nivel universitario, cuyas problemáticas y propuestas de solución han sido presentadas en [10], lo cual implicó la reestructuración del equipo de trabajo y la metodología de desarrollo. Durante el último año se avanzó en la generación del software correspondiente, que será implementado en el corriente año con el inicio del ciclo lectivo. Finalmente, a partir de la definición del Proyecto Atender, se avanzó en la integración entre ambos proyectos.

El Proyecto Atender busca, además de sus propios objetivos, la integración de las diversas actividades desarrolladas en los últimos años por parte de GEMIS, ya que entre sus actividades busca generar un componente con un motor de inferencias para procesar información de diversos orígenes basado en técnicas propias de los sistemas inteligentes; la generación de diversos componentes de procesamiento de información a través de un desarrollo basado en componentes definido en las metodologías para los proyectos del grupo [10] y la integración con software del proyecto de Intervenciones Tecnológicas.

Dentro del proyecto “La gestión del conocimiento en pequeñas y medianas fábricas de software en el AMBA”, que fuera presentado en [11], se desarrolló un trabajo final de la Especialización en Ingeniería en Sistemas de Información de UTN-FRBA, y se encuentran dos tesis de Maestría realizando sus investigaciones. Además se presentaron avances del proyecto en [12].

Finalmente, a partir de la integración entre los proyectos de Sistemas Inteligente y Tecnología Educativa se ha identificado la necesidad de evaluar el software resultante de la actividad de estos proyectos a través de un modelo de evaluación de software confiable y destinado específicamente a la

herramientas de uso educativo, por lo cual GEMIS se encuentra desarrollando un modelo de evaluación de calidad de software educativo, basado en estructuras similares a las Normas ISO 25000 y afines.

4. FORMACIÓN DE RECURSOS HUMANOS

GEMIS se encuentra conformado por investigadores formados, tesis de maestría, graduados de grado, docentes y alumnos de la carrera de Ingeniería en Sistemas de Información. Además cuenta con alumnos beneficiarios de las Becas Iniciación en Investigación y Desarrollo (BINID) de la Universidad.

Durante el presente año se prevé la presentación de avances en dos tesis de la Maestría en Ingeniería en Sistemas de Información de UTN-FRBA, la definición de un nuevo plan de tesis en el marco de dicha maestría, la presentación de dos trabajos finales de la Especialización en Ingeniería en Sistemas de Información de UTN-FRBA y el acompañamiento e integración en algunas actividades y etapas de tesis de maestría de la Universidad Tecnológica Nacional, Facultad Regional La Plata y la Facultad de Ciencias Económicas de la Universidad de Buenos Aires.

Con el inicio del ciclo lectivo en UTN-FRBA se implementará y se podrán en marcha diferentes componentes de software desarrollados en el marco de los proyectos. Estos componentes serán el registro de asistencia para notificación temprana de docentes, seguimiento de resultados de evaluaciones (a través de la obtención de datos desde diversas fuentes como planillas, tableros de Trello e informes de Google Docs) y un tablero de revisión general de situación de los alumnos. Los resultados de las

implementaciones serán presentados en el corriente año.

5. BIBLIOGRAFÍA

[1] Bara, Marc (2015). La necesidad creciente de la Gestión del Portafolio de proyectos. OBS Business School. Barcelona. España.

[2] Mathur, S. (2006). Project portfolio management techniques. Paper presented at PMI® Global Congress 2006—Asia Pacific, Bangkok, Thailand. Newtown Square, PA: Project Management Institute. Disponible en <https://www.pmi.org/learning/library/project-portfolio-management-techniques-7624>

[3] García de Fanelli, A. (2005). Universidad, organización e incentivos. Desafíos de la política de financiamiento frente a la complejidad institucional. Editorial Miño y Ávila.

[4] Vegega, C., Bazet, A., Pividori, A., Deroche, A., Ramon, H., Fabbro, V., Straccia, L., Pytel, P., Pollo-Cattaneo, M.F. (2017). Evaluación de Datos de Desempeño de Alumnos de la Asignatura ‘Sistemas y Organizaciones’ mediante Método basado en Emparrillado. Memorias del 5to Congreso Nacional de Ingeniería Informática y Sistemas de Información.

[5] Straccia, L; Deroche, A; Pytel, P; Pollo-Cattaneo, Ma F (2017). Explotación de Información para el análisis de los procesos de enseñanza. Caso de estudio: el rol del Ingeniero y las percepciones de los alumnos. En De Giusti, A; Pesado, P. (eds). XXIII Congreso Argentino de Ciencias de la Computación CACIC 2017: libro de Actas. La Plata: Universidad Nacional de La Plata. Facultad de Informática, 2017.

[6] Straccia, L.; Deroche, A.; Pytel, P.; Pollo-Cattáneo, M.F. (2018). Information Mining for the Analysis of Teaching Processes. Case Study: Perceptions of the Students about the Role of the Information System Engineer. Trabajo presentado en Journal of Computer Science & Technology. En evaluación.

[7] Straccia, Luciano; Pytel, Pablo; Vegega, Cynthia; Pollo-Cattaneo, María Florencia (2016). Proyectos educativos como solución a problemas hallados en el proceso de enseñanza y aprendizaje en una cátedra de Ingeniería en Sistemas de Información. XI Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET). Universidad de Morón, Buenos Aires.

[8] Deroche, Ariel; Acosta, Mariana; Vegega, Cinthia; Bernal, Luciano; Straccia, Luciano; Pytel, Pablo; Pollo-Cattáneo, María Florencia (2015). Diseño de Aplicación Móvil para Asignatura de grado en Ingeniería en Sistemas de Información. Congreso Nacional de Ingeniería Informática y Sistemas de Información. CONAIIISI. Universidad Tecnológica Nacional, Facultad Regional Buenos Aires.

[9] Straccia, Luciano; Marino Aguirre, Martín; Acosta, Mariana; Vegega, Cinthia; Pytel, Pablo; Pollo-Cattáneo, María Florencia (2015). El desarrollo de artefactos de tecnología informática como aporte a las intervenciones didácticas. Congreso Nacional de

Ingeniería Informática y Sistemas de Información. CONAIIISI. Universidad Tecnológica Nacional, Facultad Regional Buenos Aires.

[10] Straccia, Luciano; Pytel, Pablo; Pollo-Cattaneo, María Florencia (2016). Metodología para el desarrollo de software en proyectos de I+D en el nivel universitario basada en Scrum. Libro de Actas del XXII Congreso Argentino de Ciencias de la Computación (CACIC).

[11] Straccia, L; Maulini, A; Pytel, P; Masci, M; Vegega, C; Pollo-Cattaneo, Ma. F (2017). La Gestión del Conocimiento en Pequeñas y Medianas Fábricas de Software en el Área Metropolitana de Buenos Aires. Proceedings XIX Workshop de Investigadores de Ciencias de la Computación, WICC 2017.

[12] Maulini, A. (2017). Modelos de gestión del conocimiento y su aplicación en software factories. Revista Latinoamericana de Ingeniería de Software. Red de Ingeniería de Software de Latinoamérica. En evaluación.

Marco de trabajo para la evaluación de usabilidad de objetos de aprendizaje basados en Realidad Aumentada

Lucas G. Kucuk¹, Jorge Ierache², Gladys Dapozo³

¹ Departamento de Informática, Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones, ¹kucuk@hotmail.es

² Grupo de Realidad Aumentada, Departamento de Ingeniería e Investigaciones Tecnológicas, Universidad Nacional de la Matanza,

² Facultad de Ingeniería, Universidad de Buenos Aires, ²jierache@unlam.edu.ar

³ Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, ³gndapozo@exa.unne.edu.ar

RESUMEN

En el presente artículo se expone una línea de trabajo orientada a evaluar la usabilidad en objetos de aprendizajes basados en realidad aumentada, en la que se abordará el estudio de estándares. El mismo integra el análisis de la aplicación de realidad aumentada en objetos de aprendizaje como soporte pedagógico en escuelas técnicas de la provincia de Misiones. Esto se despliega en una línea de investigación que aborda el desarrollo de un producto software basado en realidad aumentada acorde a estándares de calidad y usabilidad.

CONTEXTO:

El proyecto se desarrolla en el marco del Trabajo Final de la Maestría en Tecnologías de la Información de la Facultad de Ciencias Exactas, Químicas y Naturales de la Universidad Nacional de Misiones.

1. INTRODUCCIÓN

Según [1] y [2], se manifiesta que la relación existente entre los objetos de aprendizaje basados en RA y la calidad del mismo, vistos como un producto software, permiten analizar si la calidad del producto indicaría mejoras en los índices de usabilidad y en el proceso de enseñanza-aprendizaje de los usuarios. A partir del análisis de

información referido al tema, se puede afirmar que el objetivo principal de esta línea es definir un marco de trabajo de usabilidad de objetos de aprendizaje basados en realidad aumentada.

1.1. Objetos de Aprendizaje

La revolución tecnológica actual requiere una revolución también en el campo educativo, una transformación más profunda en las prácticas de enseñanza y aprendizaje y no sólo en los materiales educativos. En [1] y [11] se afirma que los docentes corren el riesgo de ofrecer “vino viejo en botella nueva”. Requiere, por tanto, de “pedagogías emergentes” [12] que exploren con éxito las posibilidades que propician las nuevas tecnologías.

1.2. La realidad aumentada como innovación para la enseñanza

En la actualidad, la tecnología ha tomado un rol preponderante en cualquier ámbito. En todos los casos, el tener conocimiento sobre las herramientas tecnológicas que pueden utilizarse en cada campo resulta importante, y hasta incluso, no tenerlo puede ser excluyente. Por ello, resulta un gran reto conseguir que las nuevas propuestas tecnológicas estén al alcance de todas las personas, y a su vez, la interacción con éstas sea lo más natural posible.

En el caso concreto de la RA (desde ahora realidad aumentada) las lecciones extraídas de su aplicación educativa indican que las mejores prácticas responden a un

enfoque pedagógico de legado constructivista, orientado al aprendizaje activo (“learning by doing”), puesto que los alumnos son quienes deciden cómo combinar la información aumentada o cómo interactuar con la simulación virtual. La relación del estudiante con el objeto de aprendizaje no está basada, por tanto, sólo en la consulta de un contenido intelectual sino que implica una experiencia de inmersión en el entorno de aprendizaje [1][2].

La RA se puede definir en base al Reality-Virtuality Continuum presentado por Milgram y Kishino [7], donde se la define como la integración de elementos reales y virtuales, pero considerándola más cercana al mundo real. Es una tecnología que complementa la percepción e interacción con el mundo real y permite al usuario estar en un entorno real aumentado, con información adicional generada por la del entorno real) y la Virtualidad Aumentada (más cerca del entorno virtual).

La RA es posible gracias a diferentes tecnologías que permiten esta experiencia aumentada de la realidad. Para esto, son fundamentales tanto los dispositivos de display utilizados para que el usuario cuente con el entorno aumentado que surge de combinar el mundo real y virtual, como también las técnicas de tracking y registración que, en base a diferentes sensores y mecanismos pueden detectar la posición y orientación de un objeto para situar de manera correcta el contenido virtual sobre el mundo real. Los resultados muestran que esta tecnología impacta positivamente en los estudiantes, ya que los ayuda a vivenciar experimentos que no podrían suceder en un laboratorio real (por ejemplo, la visualización de las moléculas en movimiento).

En [6] se realiza un relevamiento de varios ejemplos de RA aplicados en distintas asignaturas tales como física, química, geografía y matemática, que incluyen además ejemplos de juegos educativos de educación primaria que pueden ser aplicados al nivel educativo secundario. Vinculan estas experiencias de RA con diferentes teorías educativas como el aprendizaje multimedia, la teoría del aprendizaje basado en experiencias y

la teoría de la visión animada. En sus conclusiones afirman que algunas características de diseño de las actividades de RA pueden reducir la carga cognitiva de los estudiantes y permitir interacciones en forma más natural para adquirir el conocimiento disciplinar correspondiente. En [13] se describe el desarrollo e implementación de un framework que se categoriza como objeto de aprendizaje planteado desde la propuesta de un juego de mesa utilizando tecnologías de realidad aumentada, lo cual también afirma haber logrado excelentes resultados desde el punto de vista pedagógico.

Las nuevas tecnologías aportan al campo de la educación aspectos innovadores que suponen una mejora cualitativa en las formas de enseñar y aprender. En particular, el aporte de la Realidad Aumentada (RA) está siendo cada vez más reconocido por los investigadores de la educación, donde se destaca que la coexistencia de objetos virtuales y entornos reales permite a los estudiantes visualizar conceptos abstractos y relaciones espaciales complejas. Esta situación se valora especialmente en el ámbito de la enseñanza de la Informática [8].

Si se analiza los objetos de aprendizaje basados en RA visto como un producto software, es necesario indicar que actualmente se encuentra un espacio intelectual poco estudiado que se refiere a la calidad del mismo, no solo considerando los aspectos técnicos de la calidad de software del objeto, sino también los aspectos consecuentes de la utilización de éstos en la educación.

1.3. Usabilidad en aplicaciones de realidad aumentada

Durante la última década, el avance logrado por los dispositivos móviles ha dado origen a una nueva generación de desarrollo de software, impulsando el crecimiento de aplicaciones específicas y estimulando la creación de un mercado asociado. Sin embargo, a pesar de este avance, la limitante es el acompañamiento de criterios para garantizar la usabilidad y la adaptación de técnicas apropiadas para su evaluación. Si bien se han realizados grandes esfuerzos por

mejorar la recepción y aceptación por parte del usuario, el estudio de la usabilidad está marcado por criterios de evaluación generales, los cuales no se adaptan a las particularidades de contextos de uso específico, como en las aplicaciones de Realidad Aumentada [5] [6].

La escasez de especificaciones que se presentan al realizar una evaluación sobre este contexto, permite investigar y definir una visión alternativa que brinde un aporte significativo en el proceso de evaluación de usabilidad. Uno de los propósitos principales de este trabajo es explotar los conceptos generales definidos en las normas y procedimientos evaluados para obtener un mayor grado de especificación y poder llevar a cabo una evaluación que dé un valor agregado a la hora de medir la usabilidad sobre las aplicaciones del contexto.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para lograr los objetivos de este proyecto se proponen las siguientes etapas y objetivos:

- **Definir un marco de trabajo para la evaluación de usabilidad de objetos de aprendizaje basado en realidad aumentada.**

Del relevamiento previo de antecedentes en el tema, se detecta una cantidad significativa de trabajos académicos que exponen el uso de tecnologías de realidad aumentada aplicadas a la educación. En las aulas, los docentes han ratificado sus resultados positivos, por ejemplo en [9] se presenta un estudio en el cual se tomó la opinión de 81 docentes, los resultados muestran que la RA puede ser tomada como una herramienta inclusiva en las aulas, salvo para las personas con discapacidad visual, además se destaca su potencial para conciliar la brecha digital, pero presentando la necesidad de que los docentes tengan conocimientos informáticos básicos para emplearla. Por su parte, en [10] se resalta que por medio de la masificación de los dispositivos móviles y con la web 2.0, los profesores poseen nuevas herramientas para motivar a los estudiantes de

diferentes niveles educativos y en diversos campos de estudio, por ejemplo, se relata en este trabajo, una experiencia educativa con una aplicación móvil llamada “AMBAR” que integra RA para el aprendizaje de idioma inglés [11]. Si bien existen muchos estudios como los mencionados, que indican una mejora en el proceso de enseñanza-aprendizaje, no se define un marco de trabajo que unifique criterios de usabilidad de los objetos de aprendizaje que intervienen.

- **Desarrollo de objetos de aprendizaje basados en realidad aumentada.**

Esta etapa integrará un profundo análisis para el desarrollo de objetos de aprendizajes que involucren la utilización de realidad aumentada y que además se ajuste al marco de trabajo o método definido. Este proceso debe involucrar los requerimientos técnicos y pedagógicos por parte del/los usuario/s que la utilicen y lo que propicie el marco de trabajo propuesto (estándares, mediciones de calidad, criterios de usabilidad, etc).

Considerando que el ambiente de aplicación del marco de trabajo a definirse en este proyecto, así como también la evaluación de usabilidad de objetos de aprendizaje basados en realidad aumentada se dará en el marco de una institución educativa secundaria técnica de la provincia de Misiones, se realizará un estudio de criterios adaptativos al contexto con el fin de determinar la eficiencia y eficacia de las actividades del marco de trabajo propuesto en el proceso de aprendizaje.

- **Validación del marco de trabajo en objetos de aprendizaje.**

Los objetos de aprendizajes serán validados desde un producto software diseñados como soporte pedagógico para docentes y alumnos de la Escuela Provincial de Educación Técnica N° 23, nivel secundario; lo cual requerirá el diseño de intervenciones en dicha institución para la articulación de contenidos curriculares y recopilación de la información

obtenida. Este objetivo integra el análisis de resultados obtenidos a partir de la validación efectuada.

La fase de validez del marco de trabajo en términos de eficiencia y eficacia, y del producto software obtenido, es evaluada desde la combinación de dos modelos de trabajo:

a) KMbM (Knowledge-based and Model-based Methods).

Estos se basan en los conocimientos disponibles y no requieren el acceso directo a los usuarios. Por lo que son especialmente útiles cuando no es posible recoger datos directamente de los usuarios. Está subdividido en tres categorías: Evaluación por expertos, evaluación basada en documentos y evaluación basada en modelos (Heuristic Evaluation)[7][13].

b) DGM (Data Gathering Methods).

Estos métodos se centran en las formas de recopilar conocimiento sobre las características relevantes de los usuarios, las tareas y el contexto en el que se utilizan los sistemas interactivos. Algunos ejemplos: entrevistas, cuestionarios o “thinking aloud”[13].

El producto software basado en realidad aumentada que se propone en este proyecto, será evaluado en términos de “calidad del producto software”. Mientras que los objetos de aprendizaje obtenidos a partir de él, serán evaluados a partir de “evaluación heurística”[5]. Este tipo de evaluación permite identificar si la facilidad de uso está ligada a la satisfacción del usuario. También permite analizar que la consecución rápida de objetivos está ligada a la efectividad, para este caso, se refiere a objetivos de aprendizaje. Finalmente, la utilización de recursos es una característica que este tipo de evaluación permite evaluar para obtener la eficiencia.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Si bien el proyecto se encuentra en sus etapas iniciales, actualmente se ha realizado un relevamiento bibliográfico respecto a modelos de usabilidad utilizados y aplicables en objetos de aprendizaje basados en realidad aumentada.

Así también, se encuentra en desarrollo una aplicación como objeto de aprendizaje basado en RA, bajo plataforma Vuforia Developer Portal en conjunto con Android Studio, teniendo en cuenta los estudios anteriormente citados y los métodos DGM (Data Gathering Methods) [3] y KMbM (Knowledge-based and Model-based Methods)[6][13].

La contribución más novedosa es que el marco de trabajo permite el intercambio de roles entre usuarios adaptando la usabilidad del producto a la necesidad de aprendizaje.

El resultado esperado es formalizar el marco de trabajo mediante el desarrollo de la aplicación basada en realidad aumentada que posibilite el acceso a objetos de aprendizaje, además de la validación de la usabilidad con usuarios (docentes y estudiantes) pertenecientes a una escuela de formación técnica de nivel secundario de la provincia de Misiones.

La validación, pruebas de ejecución y recopilación de información sobre los resultados de la usabilidad de los objetos de aprendizajes basados en realidad aumentada y aplicación del marco de trabajo definido se realizará en la Escuela Provincial de Educación Técnica N° 23 de la provincia de Misiones, con alumnos y docentes desde el primer año hasta el sexto con orientación en electromecánica. Se ha seleccionado esta escuela, debido a que esta propuesta puede extenderse a toda la formación técnica de la provincia y es un área sin precedentes de aplicación de esta tecnología.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de trabajo participan los directores, docentes investigadores de la Universidad Nacional de La Matanza y la Universidad Nacional del Nordeste, con antecedentes de investigación en los temas que

se abordan. El desarrollo del proyecto será realizado por un estudiante de la Maestría en Tecnologías de la Información de la Facultad de Ciencias Exactas, Químicas y Naturales de la Universidad Nacional de Misiones.

5. BIBLIOGRAFÍA

- [1] Jorge Martín-Gutiérrez, José Luís Saorín, Manuel Contero, Mariano Alcañiz, David C. Pérez-López, y Mario Ortega. Design and validation of an augmented book for spatial abilities development in engineering students. *Computers & Graphics*, 34(1), pp.77-91, 2010. ISSN 0097- 8493.
- [2] Ángela Di Serio, María Blanca Ibáñez, y Carlos Delgado Kloos. Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education*, 68(0), 586-596, 2013. ISSN 0360-1315.
- [3] T.G. Kirner, F.M.V. Reis, y C. Kirner. Development of an interactive book with augmented reality for teaching and learning geometric shapes. En 7th Iberian Conference on Information Systems and Technologies (CISTI 2012), pp. 1-6, 2012.
- [4] Natali Salazar Mesía, Cecilia Sanz, Gladys Gorga, Experiencia de enseñanza de programación con Realidad Aumentada, Actas de las XXII Jenui. Almería, 6-8 de julio 2016, Universidad Nacional de La Plata, La Plata, Buenos Aires, Argentina ISBN: 978-84-16642-30-4 Páginas: 213-220.
- [5] MEDINA, Guillermo Pablo; GIRARDI, Hernán. Análisis de usabilidad de aplicaciones de realidad aumentada en dispositivos móviles: un procedimiento para la medición y evaluación. 2013. Tesis Doctoral. Facultad de Informática.
- [6] Marc Ericson C. Santos, Angie Chen y Takafumi Taketomi. Augmented Reality Learning Experiences: Survey of Prototype Design and Evaluation 2014. *IEEE Transactions On Learning Technologies*, vol. 7, no. 1, pp. 38-56, marzo de 2014.
- [7] Paul Milgram, Haruo Takemura, Akira Utsumi, Fumio Kishino. Augmented Reality: A class of displays on the reality-virtuality continuum. 1994. *Telem manipulator and Telepresence Technologies*, 2351, pp. 282-292.
- [8] Rana M. Tamim, Robert M. Bernard, Eugene Borokhovski, Philip C. Abrami y Richard F. Schmid. What Forty Years of Research Says about the Impact o Technology on Learning a Second-Order Meta-Analysis and Validation Study. *Review of Educational Research*, marzo de 2011, Vol. 81, No. 1, pp. 4-28.
- [9] V. Díaz Marín, “Posibilidades de uso de la Realidad Aumentada en la educación inclusiva. Estudio de caso,” *ENSAYOS. Rev. la Fac. Educ. Albacete*, vol. 31, no. 2, pp. 57–67, 2016.
- [10] M. Morales Hernández, C. Benitez Quecha, D. Silva Martinez, M. Altamirano Cabrera, and H. Mendoza Gómez, “Aplicación móvil para el aprendizaje del inglés utilizando realidad aumentada,” *Rev. Iberoam. Prod. Académica y Gestión Educ.*, 2015
- [11] KNOBEL, Michele; LANKSHEAR, Colin (ed.). *DIY media: Creating, sharing and learning with new technologies*. Peter Lang, 2010.
- [12] ADELL, Jordi; CASTAÑEDA, Linda. Tecnologías emergentes,¿ pedagogías emergentes. *Tendencias emergentes en educación con TIC*, 2012, p. 13-32.
- [13] Ierache J., Mangiarua N., Bevacqua S., Becerra M., Verdicchio N., Duarte N., Sanz D., Ortiz F, Igarza S. 2014. “Herramienta de Realidad Aumentada para facilitar la enseñanza en contextos educativos mediante el uso de las TICs”. *Revista Latinoamericana de Ingeniería de Software*, 1(1): 1-3, ISSN 2314-264

Computación Aplicada en Ciencias y Educación

Adair Martins, Carina Fracchia, Claudia Allan, Susana Parra,
Natalia Baeza, Carolina Celeste, Nahuel Mamani, Kevin Isaias Pascual,
Ana Alonso de Armiño, Roberto Laurent

Departamento de Computación Aplicada / Facultad de Informática
Departamento de Electrotecnia / Facultad de Ingeniería
Universidad Nacional del Comahue

Dirección: Buenos Aires 1400, 8300 – Neuquén

Teléfono: 0299 - 4490300 int. 429

e-mails: {adair.martins, carina.fracchia, claudia.allan, susana.parra }@fi.uncoma.edu.ar,
{baeza.natalia, celeste.carolina.s, anaalonso}@gmail.com,
{nahuel.mamani, kevin.isaias}@est.fi.uncoma.edu.ar, {roberto_laurent}@yahoo.com.ar

Resumen

La utilización de nuevas herramientas tecnológicas en el ámbito educativo es un factor estratégico clave para la mejora y fortalecimiento de los procesos de enseñanza y aprendizaje en los distintos campos del conocimiento, principalmente en las carreras de ciencias e ingeniería. En base a este contexto se presentan las líneas principales en las que se viene trabajando: Desarrollo y Uso de Recursos TIC, Realidad Aumentada y Realidad Virtual, Métodos Computacionales y Simulación, Objetos de Aprendizaje. Las mismas están interrelacionadas, persiguiendo como objetivo general acercar la Universidad al medio. En este trabajo se presentan los fundamentos y avances de la investigación que se está realizando.

Palabras clave: Computación Aplicada, Recursos TIC, Educación, Métodos Computacionales, Realidad Aumentada, Objetos de Aprendizaje

Contexto

Las líneas de investigación y desarrollo se enmarcan en el Proyecto de Investigación “Computación Aplicada en Ciencias y Educación”, Facultad de Informática (FAIF), Universidad Nacional del Comahue (UNCo). Está integrado por docentes y estudiantes de

la carrera de Licenciatura en Ciencias de la Computación de la FAIF y docentes de la Facultad de Ingeniería, UNCo, y de la Universidad Católica de Brasilia (UCB), Brasil.

Introducción

En la asignatura Métodos Computacionales para el Cálculo de las carreras de Licenciatura en Ciencias de la Computación y Profesorado en Informática de la Facultad de Informática se dictan contenidos del cálculo diferencial e integral en varias variables con aplicaciones de métodos numéricos. Uno de los objetivos de la materia es familiarizar al estudiante con el uso de herramientas computacionales como apoyo en la resolución de problemas matemáticos y numéricos. En trabajos referenciados en la bibliografía se describen las dificultades encontradas por los estudiantes en la comprensión de los conceptos teóricos y prácticos del cálculo diferencial e integral en una o más variables, de las funciones multivariantes y en la interpretación geométrica de los métodos numéricos. En las últimas décadas fueron desarrollados varios sistemas de cálculo simbólico que pueden ser clasificados como Sistemas de Computación Algebraica (CAS) según sus características, y que en general trabajan con formas algebraicas, numéricas y

gráficas. Se observaron limitaciones en los software utilizados en el dictado de la asignatura mencionada en cuanto a la interfaz y a la sintaxis para realizar los cálculos en los métodos numéricos y en los métodos del cálculo diferencial [1-4]. De lo mencionado anteriormente se propuso el desarrollo de una herramienta online basada en software libre como apoyo a la enseñanza de los métodos computacionales. La misma forma parte del trabajo final de la carrera de Licenciatura en Ciencias de la Computación de una integrante alumna del Proyecto de Investigación.

Continuando con el estudio de los métodos computacionales, se puede mencionar que en la gráfica de funciones multivariables el concepto de visualización adquiere un rol fundamental para la comprensión de los contenidos en la enseñanza del Cálculo diferencial e integral en multivariables. Es de gran importancia explorar las herramientas de visualización que ofrecen las nuevas tecnologías, posibilitando nuevos ambientes de enseñanza-aprendizaje [5]. En esta línea se continúa con el trabajo realizado en investigaciones anteriores en Objetos de Aprendizaje (OA), estudiando el potencial de las herramientas de geometría dinámica tridimensional para desarrollar OA que faciliten la práctica docente y la comprensión de los temas del cálculo multivariado. El objetivo es la implementación de nuevos objetos poniendo foco en las tecnologías de visualización tridimensional y su acceso desde distintos dispositivos de comunicación. La internet supone la vía de desarrollo del aprendizaje virtual que, apoyado en plataformas de aprendizaje, posibilitará una optimización del proceso formativo. Gracias a la difusión de los dispositivos móviles tales como teléfonos celulares, tablets, etc., se está transitando a una nueva etapa del proceso, el m-Learning [6-7]. Se sigue trabajando en la adaptación y estandarización de los recursos educativos desarrollados para facilitar su visualización a través de dispositivos móviles. Principalmente apuntando a un diseño accesible para ser utilizado en asignaturas

dictadas en las carreras de la Facultad de Informática.

Se sigue avanzando en la línea “Realidad Aumentada y Virtual”. Los medios por los cuales las personas interactúan con las computadoras evolucionan rápidamente, posibilitando que surjan nuevas formas de comunicación e interacción con la computadora. Un ejemplo de ello son las interfaces de Realidad Aumentada (RA) que permiten generar un entorno combinado para el usuario, por un lado información del mundo real, y por otro, información sintética creada y manejada por la computadora. La RA permite la fusión de datos virtuales y objetos sobre el mundo físico, enriqueciendo la percepción de la realidad, aumentándola. A diferencia de la Realidad Virtual (RV) no implica la inmersión en un ambiente completamente virtual, dado que se mantiene en el mundo real mientras se despliegan y muestran contenidos virtuales superpuestos sobre él [8]. Los resultados obtenidos desde el año 2015 en diferentes Proyectos de Extensión llevados a cabo en Instituciones Educativas del nivel primario y medio de la provincia han demostrado el potencial de la RA en la enseñanza. También ha demostrado su función pedagógica en otro tipo de escenarios como son los museos, estos vistos como un espacio educativo. La RA facilitará que sus visitantes no actúen como meros observadores pasivos sino que facilitará su interacción con los objetos tales como cuadros, esculturas, fósiles entre otros [9]. En la investigación que viene desarrollando el grupo de RA pretende hacer accesible las tecnologías de RA y RV a personas sin conocimientos técnicos, contribuyendo de esta manera a enriquecer sus proyectos con contenido aumentado.

Continuando con las líneas de investigación, se puede mencionar que con el avance y crecimiento de la utilización de las computadoras para procesar la información y la inmensa cantidad de páginas existentes en la web, a medida que crece internet, para satisfacer las necesidades de millones de

usuarios también aumentan los problemas relacionados con su estabilidad. En este sentido cuando un usuario utiliza la web para buscar una información en una página este procedimiento es realizado con un motor de búsqueda. Los primeros mecanismos de búsqueda en la internet se basaban en las informaciones de sus páginas, pero con el crecimiento en forma exponencial de las mismas, este se tornó insuficiente para contener las páginas con el mismo contenido. En ese momento fue creado el ranking de las páginas web denominado PageRank (PR) [10]. PR es un método que fue implementado a través de un algoritmo y que asigna un puntaje numérico para la relevancia de las páginas web indexadas por el motor de búsqueda de Google siguiendo enlaces. El PR es un súper conjunto de fórmulas matemáticas y tiene el objetivo de medir la relevancia de determinadas páginas para que puedan ser posicionadas en el motor de búsqueda de Google. Posee muchas aplicaciones prácticas por ejemplo en el cálculo del ranking en deporte, turismo, sistemas educacionales, entre otros. En función de lo expuesto anteriormente se está estudiando el funcionamiento del algoritmo de PageRank con el objetivo de identificar y analizar los métodos matemáticos y numéricos. Particularmente los métodos de resolución de ecuaciones lineales, teoría de grafos, métodos de cálculo de autovalores y autovectores, que son utilizados en su cálculo a través de ejemplos y utilización de simulación [11-12].

Líneas de Investigación y Desarrollo

En forma resumida se presentan las líneas de Investigación y Desarrollo en las que se está trabajando:

- Desarrollo y Uso de Recursos TIC
- Realidad Aumentada y Realidad Virtual.
- Objetos de Aprendizaje
- Métodos Computacionales y simulación

Resultados Obtenidos/Esperados

En la línea “Desarrollo y Uso de Recursos TIC” se encuentra en desarrollo una herramienta online basada en software libre como apoyo a la enseñanza de los métodos computacionales, poniendo énfasis en el diseño de una interfaz de usuario intuitiva y amigable que permita integrar los contenidos principales del cálculo diferencial e integral con los métodos numéricos. Cabe mencionar que este desarrollo forma parte de plan de tesis de una alumna de la carrera Licenciatura en Ciencias de la Computación. Para el desarrollo de la herramienta se pretende evaluar diversos entornos y lenguajes de programación existentes y analizar si estos poseen librerías enfocadas en resolver problemas de matemáticas y computación científica. En particular aquellas que permitan el manejo de matemática simbólica, la manipulación de matrices n-dimensionales, y librerías para gráficos en dos y tres dimensiones.

Se encuentra también en desarrollo una tesis de doctorado titulada “Propuestas de torneos en los niveles medio y universitario en el ámbito de la enseñanza de la programación: análisis de su impacto en la motivación y el rendimiento académico de los estudiantes”. Persigue como objetivo principal generar contribuciones teóricas y metodológicas relacionadas con el uso de torneos de Programación como recurso didáctico para promover una motivación positiva y mejorar el rendimiento académico en esta área. Para ello el estudio se centra en identificar dificultades que se presentan en los estudiantes respecto al aprendizaje de la programación. Además de estudiar potenciales impactos de la participación en torneos en la motivación de los estudiantes por los temas de programación. En esta línea, se propone seguir investigando sobre uso y desarrollo de recursos TIC que puedan ser utilizados en la resolución de problemas aplicados a la educación y al medio.

Continuando con el trabajo sobre Objetos de Aprendizaje se vienen explorando las potencialidades de distintos software de geometría dinámica en el espacio, poniendo énfasis en aquellos libres, de código abierto y disponibles para diversas plataformas, inclusive dispositivos móviles. Se propone implementar distintos escenarios (scripts), utilizando lenguajes de programación web para la generación de interfaces que permitan la animación y simulación del comportamiento de distintos conceptos matemáticos trabajados, permitiendo la interacción del usuario a través de la manipulación de diversos parámetros. En este sentido se están evaluando distintas metodologías de diseño de OA con el objetivo de permitir la integración con un software de geometría dinámica y su incorporación a entornos virtuales de enseñanza y aprendizaje, contemplando los distintos formatos que propicien dicha integración y el acceso tanto desde PC como desde dispositivos móviles.

Se han diseñado OA que implican la resolución de ejercicios online en diversos temas de las asignaturas de las carreras de la FAIF. Al incluirlos en el Entorno Virtual de Aprendizaje se ha favorecido tanto la autoevaluación por parte de los estudiantes como la obtención por parte de los docentes de diversas estadísticas relacionadas con la resolución que ofrece el paquete Moodle. Utilizando el software GeoGebra se diseñó y se implementó un OA para el método numérico de Newton Raphson para el cálculo de raíces y un OA para el cálculo de integrales por el método de los trapecios. El objetivo principal es permitir la visualización por parte de los estudiantes del comportamiento gráfico de estos métodos, logrando así que puedan realizar una correcta interpretación geométrica de los mismos. Se pretende por ejemplo para el método de los trapecios que a través de la manipulación del parámetro que determina la cantidad de particiones del intervalo, los estudiantes puedan comprender cómo, a medida que se aumenta ese parámetro, disminuye el error

cometido en el cálculo de la aproximación a una integral definida. Estos estudios dieron origen al trabajo publicado en la revista TE&ET: “Objetos de Aprendizaje para la Interpretación Geométrica de Métodos Numéricos: Uso de GeoGebra” [13].

En la línea de Realidad aumentada y Virtual se continúa trabajando con las escuelas del nivel primario de la Provincia de Neuquén y Río Negro a través de la realización de experiencias, talleres y cursos. En esta línea dio origen a la publicación: “Realidad Aumentada en la Enseñanza Primaria: Diseño de Juegos de Mesa para las Áreas Ciencias Sociales y Matemáticas, III Congreso Internacional de Investigación y Docencia de la Creatividad [14].

También fue desarrollado por alumnos de la Licenciatura en Ciencias de la Computación e integrantes del grupo de trabajo de Realidad Aumentada un prototipo de software para el trabajo con temas de la provincia de Neuquén. Mediante diferentes imágenes (mapa físico-político de Neuquén, Escudo y Bandera, entre otros) se activa la información referente al elemento seleccionado utilizando dispositivos como tales como celular, tablet, notebook y PC. Esta aplicación permite a través de diferentes íconos sobre el mapa acceder a la información turística de la provincia como por ejemplo el Volcán Lanín, las Termas de Copahue, entre otros. El mismo fue presentado en la DevFest Neuquén 2017 (Conferencia de Desarrolladores de todo el día), Neuquén, UNCo, 2017 [15].

En nuestra región existen varios museos dedicados al rescate, conservación, investigación y difusión del patrimonio cultural. El uso de las TIC en el campo del Patrimonio Cultural representa actualmente una novedosa herramienta para su conocimiento, permitiendo desarrollar experiencias y recursos para la difusión del objeto cultural. El diseño de recursos aumentados posibilita “aumentar” los contenidos de un determinado objeto cultural mostrando elementos del pasado que ya han desaparecido o que se corresponden con

momentos históricos diferentes, sobre la imagen real del objeto en la realidad, combinando ambas instancias históricas. El objetivo principal del estudio que se está realizando es analizar y establecer los elementos y criterios necesarios para el diseño de un prototipo de una aplicación de realidad aumentada, que pueda utilizarse en diferentes dispositivos móviles de uso personal y pensado para ser usado en un entorno de museos. Se continuará trabajando en el desarrollo de recursos RA y RV para ser aplicados en temas relacionados con área cultural, turismo y en educación.

Formación de Recursos Humanos

El grupo de investigación viene trabajando y formándose en proyectos anteriores a través de la realización de cursos de postgrado, capacitación, entre otros. Actualmente una integrante docente se encuentra en etapa de tesis en el doctorado: “Enseñanza de las Ciencias Exactas y Naturales” en la Facultad de Ingeniería (FI), UNCo. Otra integrante está finalizando el cursado de la maestría en Tecnología Aplicada en Educación en la Universidad Nacional de la Plata. Dos alumnas del grupo están finalizando su tesis de grado.

Referencias

- [1] Oye, N. D., Z.K. Shallsuku, A.Iahad, N., The Role of ICT in Education: Focus on University Undergraduates taking Mathematics as a Course, International Journal of Advanced Computer Science and Applications (IJACSA), vol. 3, N°2, 2014.
- [2] Cavasotto, M, Reflexões sobre as Dificuldades na aprendizagem de cálculo diferencial e integral, III Mostra de Pesquisa da pós-graduação, PUCRS, 2010.
- [3] Krishnachandran V.N., Reji C. Joy, K. B. SIPI, More CAS's in maths classrooms: An urgent imperative, Thechnolgy Enhanced Education (ICTEE), IEEE Internacional Conference on Education, doi: 10.1109/ICTEE.2012.6208645, 2012.
- [4] Sharma J., Kumar S. R., Strategies for Web Application Development Methodologies, International Conference on Computing, Communication and Automation (ICCCA), 2016.
- [5] Lemke, R., Objetos de aprendizagem para o ensino de funções de duas variáveis: um diferencial dinâmico, 2015.
- [6] Henriques, A., Attie, Juao P., F., Santos L. M., Referências teóricas da didática francesa: análise didática visando o estudo de integrais múltiplas com auxílio do Maple, 2007.
- [7] Aguilar, G., Chirino, V., NERI, L. Noguez, J., Tella R., Tella V., Impacto de los recursos móviles en el aprendizaje, 9ª Conferencia Iberoamericana en Sistemas, ISBN: 13: 978-1-934272-940-7), Florida, USA, 2010.
- [8] C. Manresa Yee, M.J. Abásolo, R. Más Sansó, M. Vénere. Realidad Virtual y Realidad Aumentada. Interfaces avanzadas, 1ª ed. Universidad Nacional de La Plata, La Plata. ISBN 978-950-34-0765-3, 2011.
- [9] Ruiz Torres D., Realidad aumentada y Patrimonio Cultural: nuevas perspectivas para el conocimiento y la difusión del objeto cultural. e-rph: Revista Electrónica de Patrimonio Histórico [en línea], 2011.
<http://www.revistadepatrimonio.es/revistas/numero8/difusion/estudios2/articulo.php>
- [10] Assis Machado. C., Macedo Rocha, C., Magalhães, E. J., Pereira e Silva, M. D., Utilização do Algoritmo PageRank para Obtenção de Qualidade no Ranqueamento de Páginas Web, Simpósio de Sistemas de Informação, UFG, 2016.
- [11] Bekhout J., Google's PageRank Algorithm for Ranking Nodes in General Networks, Proceedings of the 13th International Workshop on Discrete Event Systems, Xi'an, China, May 30, June 1, 2016.
- [12] Infolab Stanford
<http://infolab.stanford.edu/~backrub/google.html>
- [13] Allan C, Parra S., A. Martins, Objetos de Aprendizaje para la Interpretación Geométrica de Métodos Numéricos: Uso de GeoGebra”, TE&ET N°20, ISSN:1850-9959| RedUNCI-UNLP, 2017.
- [14] Fracchia C., A. Martins, Realidad Aumentada en la Enseñanza Primaria: Diseño de Juegos de Mesa para las Áreas Ciencias Sociales y Matemáticas, III Congreso Internacional de Investigación y Docencia de la Creatividad. Libro de resúmenes. ISBN: 978-84-697-5239-5, Universidad de Granada, España, 2017.
- [15] DevFest Neuquén 2017, UNCO.
https://www.meetup.com/es/gdneuquen/events/243356395/?_cookie-check=gBESweSv6GCJHepm

Realidades Alternativas como Herramientas de Mediación para el Desarrollo del Pensamiento Computacional

M. V. Rosas, M. Zuñiga, J. Fernández, R. Guerrero

Laboratorio de Computación Gráfica - Dpto. de Informática - FCFMyN
Universidad Nacional de San Luis
Ejército de los Andes 950 (San Luis)
{mvrosas, mezuniga, jmfer, rag}@unsl.edu.ar

Resumen

El crecimiento exponencial de las nuevas tecnologías prácticamente ha impactado en todos los ámbitos de nuestras vidas. La Realidad Virtual y la Realidad Aumentada permiten alterar la percepción de un mundo canónico, generando en consecuencia, Realidades Alternativas. En los últimos tiempos se ha evidenciado un creciente interés en incorporar la Visualización Digital Avanzada en programas educativos de todos los niveles. En este sentido, es preciso destacar que cualquiera sea la herramienta que se considere válida de ser aplicada en educación debería estar direccionada hacia el logro de aprendizajes significativos. El Pensamiento Computacional, valiéndose de estas realidades alternativas, permitiría beneficiar los procesos de enseñanza y aprendizaje involucrados en su desarrollo. De esta manera, la intención de esta investigación está centrada en estudiar las potencialidades que estas realidades alternativas aportarían al desarrollo del Pensamiento Computacional. Así mismo, estudiar y analizar las potencialidades de las pedagogías emergentes de este contexto que favorecerían la formación de los alumnos para la actual sociedad de la información en la que están inmersos.

Palabras clave: Realidad Aumentada, Realidad Virtual, Pensamiento Computacional, Pedagogías emergentes.

Contexto

Esta propuesta surge como producto de las acciones de investigación y producción desarrolladas por los integrantes del equipo en estos últimos años a través de su participación en dos proyectos nacionales (“Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos” y “Estrategias para la Mejora de la Enseñanza de la Programación a Alumnos Ingresantes de las carreras de Ciencias e Ingeniería”), y el Proyecto internacional Alfa GAVIOTA (Grupos Académicos para la Visualización Orientada por Tecnologías Apropriadas). Se pretende continuar con las temáticas abordadas en los proyectos mencionados, reestructurando el objeto de estudio con la intención de promover la sinergia entre los conceptos involucrados en el marco de un nuevo proyecto denominado “Realidades Alternativas como Lenguaje Generativo aplicado a la solución de Problemas Reales”.

1. Introducción

Las tecnologías emergentes presentan a nivel mundial un marcado crecimiento en relación a su uso. El área de investigación sobre Visualización Digital Avanzada, conformado por la Realidad Virtual (RV) y la Realidad Aumentada (RA), ha contribuido a incrementar el uso de dichas tecnologías y ampliado sus campos de aplicación.

Las RV sumerge al usuario en un ambiente generado por computadora que simula la realidad mediante el uso de dispositivos interactivos que envían y reciben información

(gafas, cascos, guantes, trajes). De esta manera el usuario puede realizar recorridos virtuales de ambientes 3D simulados, al mismo tiempo que experimenta la sensación de tocar, capturar y manipular los objetos que está visualizando. En términos generales, la RV se constituye en una compleja interfaz de usuario que engloba simulaciones e interacciones en tiempo real a través de múltiples canales sensoriales. Estas modalidades sensoriales son visuales, auditivas, táctiles, olfativas, kinestésica, entre otras. En la actualidad, si bien el campo de aplicación de la RV ha crecido significativamente, produciendo una enorme expansión de esta tecnología, queda un camino muy interesante por recorrer donde la misma pueda ser utilizada en una generalidad de situaciones [2, 3].

Con respecto a la RA, ésta supone la introducción de elementos virtuales en el mundo real. Se trata de generar objetos, seres, contextos, imágenes y textos virtuales, entre otros, por medio del ordenador, que puedan superponerse o “incluirse” en el mundo real. Esto es, el usuario está viendo el contexto real en el que se encuentra por medio de una o más cámaras y, a la vez, está viendo esos elementos virtuales. La característica central de las aplicaciones de RA es que los elementos virtuales superpuestos en el mundo real proporcionan información adicional y relevante a la imagen final que está visualizando el usuario del sistema con el objetivo de ayudarlo. Finalmente, si bien la RA ya ha demostrado su utilidad en algunas áreas (entrenamiento, ingeniería, juegos y entretenimientos, diseño) y es reciente en otras (salud y educación), los logros obtenidos surgen de aplicaciones limitadas donde las expectativas se centran en la maduración de estas tecnologías [4, 5].

El campo educativo no ha quedado fuera de este progresivo interés en el uso de las realidades alternativas. Programas educativos de todos los niveles las han incorporado como herramienta para el enriquecimiento de la transmisión de diferentes conceptos. Puntualmente, el uso de la RV en el ámbito

educativo ha realizado notables aportes en áreas como la Biología, Historia y Geografía mediante la simulación y recreación de escenarios complejos de vivenciar en la realidad, brindando a los estudiantes la posibilidad de participar en experiencias inmersivas [6, 7, 8].

Por otro lado, en 2006 Jeannette Wing introduce al Pensamiento Computacional (PC) como la manera de resolver problemas y entender el comportamiento humano valiéndose del aprendizaje de conceptos básicos de las ciencias de la computación a través de los pilares fundamentales del PC. “*El Pensamiento Computacional consiste en la resolución de problemas, desarrollando habilidades asociadas a conceptos fundamentales de la informática...*”. Wing asegura que el pensamiento computacional inevitablemente está invadiendo las demás disciplinas. No sólo científicos de la computación hacen uso de este tipo de pensamiento. Es una habilidad y una actitud de aplicación universal para todas las personas, por lo que el PC beneficiará a los estudiantes y a las instituciones en todos los niveles educativos, contribuyendo en el desarrollo humano, científico y tecnológico [9,10]. De esta manera, el PC beneficiaría a una amplia variedad de disciplinas, al mismo tiempo que él evolucionaría en la medida que más se lo utilice en una diversidad de áreas temáticas.

Diversos autores argumentan que la incorporación del PC en las disciplinas tiende a afianzar el conocimiento e innovar en sus prácticas apartándose de esquemas habituales y abordando la realidad de manera novedosa [11, 12, 13].

Actualmente, la educación está siendo atravesada por diferentes cambios y reestructuraciones. Una ruptura destacada se observa sobre las estrategias convencionales de enseñanza a partir de las nuevas concepciones del aprendizaje, que consideran al sujeto como eje central en los procesos de enseñanza y aprendizaje. Esto último provoca un cambio metodológico y establece un desafío para la enseñanza. Los docentes deben

aprender a comunicarse con los nuevos lenguajes y acorde a los nuevos estilos de aprendizaje, repensar el lugar de los contenidos a enseñar, lo que resultaría en un desplazamiento desde el paradigma basado en el conocimiento al paradigma basado en la acción e interacción [14, 15, 16].

A partir de lo expuesto, es posible advertir la pertinencia de integrar herramientas de Visualización Digital Avanzada en estos nuevos modelos educativos. Pero, es importante destacar que para que dicha integración sea beneficiosa debería estar direccionada hacia el logro de aprendizajes significativos. El desarrollo del PC podría permitir delimitar esta direccionalidad. Así, el PC asistido por las realidades alternativas no solo presenta un alto potencial para mejorar la eficiencia de los procesos de enseñanza y aprendizaje sino que, además, puede actuar como innovación disruptiva que promueva un nuevo enfoque pedagógico en el marco de las llamadas pedagogías emergentes [17, 18].

La integración de los conceptos abordados en este trabajo es, hasta donde conocemos, una propuesta novedosa en el contexto de la investigación nacional e internacional, abriendo un espacio de exploración rico en líneas de abordaje.

2. Líneas de Investigación y Desarrollo

Los ejes que delimitan esta investigación están enfocados en explorar las sinergias de un ecosistema de aprendizaje que involucre elementos innovadores y disruptivos resultantes del uso de tecnologías como RA y RV. En consecuencia, se plantean tres ejes:

Eje 1: La RA y la RV como herramientas para potenciar el desarrollo del PC y consolidar sus pilares fundamentales.

Eje 2: La Realidad Mixta como herramienta para potenciar el desarrollo del PC y consolidar sus pilares fundamentales.

Eje 3: Las pedagogías emergentes que surgen alrededor del uso de estas tecnologías emergentes.

Se ha realizado un valioso avance sobre los objetivos planteados inicialmente (Eje 1) a la vez que se pretende continuar con los logros obtenidos, extendiendo el objeto de estudio en forma más focalizada a la comprensión de las interrelaciones existentes entre las tecnologías mencionadas y aquellas áreas con mayor factibilidad de aplicación, entre las cuales se pueden destacar al área de la salud y la educación (Eje 2 y 3).

3. Resultados obtenidos/esperados

En consonancia con los ejes planteados se pretende:

- Descubrir los factores que inciden en el desarrollo de estrategias didácticas para mejorar la formación de docentes en el uso de las realidades alternativas.
- Generar conocimiento sobre las pedagogías emergentes, con la intención de contribuir a la comprensión y el análisis de su impacto y potencial disruptivo en el campo educativo.
- Promover espacios de formación académica en relación a las temáticas planteadas con la intención de contribuir a la capacitación docente.
- Fomentar en las prácticas educativas el desarrollo del PC mediado por las realidades alternativas para la resolución de problemas generales favoreciendo el aprendizaje significativo.

4. Formación de Recursos Humanos

En el equipo de trabajo participan docentes con formación de grado y posgrado en carreras relacionadas a la Informática y a la Educación Superior.

Los trabajos realizados hasta el momento han permitido la concreción de trabajos de fin de carrera de la Licenciatura en Ciencias de la Computación y una (1) Especialización en Educación Superior, así como también la definición de una (1) tesis de Especialización

y dos (2) trabajos de tesis de Maestría en Educación Superior.

5. Bibliografía

- [1] Aukstakalnis S.. “Practical Augmented Reality: A Guide to the Technologies, Applications, and Human Factors for AR and VR”. Usability Series. Addison Wesley Professional, 2016.
- [2] Vince J.. “Introduction to Virtual Reality”. Springer London, ISBN 978-0-85729-386-2, 2011.
- [3] Steinicke F.. “Being Really Virtual: Immersive Natives and the Future of Virtual Reality”. Springer International Publishing, 2016.
- [4] Serón Arbeloa F. J., Carvalho C., Saul M., Deinzer F., Guerrero R., Parra Márquez J. et al., “Realidad virtual, realidad aumentada y computación ubicua: conceptos, tecnología y práctica”. Realidades virtuales aumentadas para el desarrollo social: Experiencias entre Europa y Latinoamérica, pp. 47-68, ISBN 978-950-757-046-9, Editor: Mónica Inés Fernández, Editorial Universidad de Belgrano, Buenos Aires, 2014.
- [5] Peddie J.. “Augmented Reality Where We Will All Live”. Springer International Publishing eBook ISBN 978-3-319-54502-8 DOI 10.1007/978-3-319-54502-8 Hardcover ISBN 978-3-319-54501-1, 2017.
- [6] Chen P., Liu X., Cheng W., y Huang R. “A review of using Augmented Reality in Education from 2011 to 2016”. En E. Popescu, Kinshuk, M. K. Khribi, R. Huang, M. Jemni, N.-S. Chen, y D. G. Sampson (Eds.), Innovations in Smart Learning. (13-18). Singapore: Springer, 2017.
- [7] Yuen, S. Chi-Yin; Yaoyuneyong, G.; y Johnson, E.. “Augmented Reality: An Overview and Five Directions for AR in Education” Journal of Educational Technology Development and Exchange (JETDE): Vol. 4 : Iss. 1 , Article 11. DOI: 10.18785/jetde.0401.10 Available at: <https://aquila.usm.edu/jetde/vol4/iss1/11>, 2011.
- [8] Aukstakalnis S.. “Practical Augmented Reality: A Guide to the Technologies, Applications, and Human Factors for AR and VR”. Usability Series. Addison Wesley Professional, 2016.
- [9] Wing J. M.. “Computational Thinking and Thinking about Computing”, Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, p p. 3717-3725, 2008.
- [10] Wing J.M.. “Computational thinking”, Commun. ACM, p. 49(3):33–35, 2006.
- [11] Cuny, J., Snyder, L. y Wing, J.M. “Demystifying computational thinking for non computer scientists”. Unpublished manuscript in progress, referenced in <http://www.cs.cmu.edu/~CompThink/resources/TheLinkWing.pdf>. (2010)..
- [12] Linn, M., Aho, A., Blake, B., et al.: “Report of a Workshop on the pedagogical aspects of Computational Thinking”. The National Academies Press. ISBN 978-0309-21474-2. (2011).
- [13] Linn, M., Aho, A., Blake, B., y Constable, R.. “Report of a Workshop on the scope and nature of Computational Thinking.” Whashington, D.C.: The National Academies Press. (2010).
- [14] Gisbert, M.. “Digital Leaners: la competencia digital de los estudiantes universitarios.” La Cuestión Universitaria, 48-59. (2011).
- [15] Stager, G.. “En pro de los computadores”, 13 enero 2004. [En línea]. Available: <http://www.eduteka.org/ProComputadores.php>.
- [16] Adell, J. “Los estudiantes universitarios en la era digital: la visión del profesor”. La Cuestión Universitaria, pp 97-100, 2011.
- [17] Adell, J. y Castañeda, L. Tecnologías emergentes, ¿pedagogías emergentes? En

- J. Hernández, M. Pennesi, D. Sobrino y A. Vázquez (coord.). Tendencias emergentes en educación con TIC. Barcelona: Asociación Espiral, Educación y Tecnología. págs. 13-32. ISBN: 978-84-616-0448-7. (2012).
- [18] Veletsianos, G. "A definition of emerging technologies for education". En Veletsianos, G. (ed.) Emerging technologies in distance education (pp. 3-22). Athabasca, CA: Athabasca University Press, 2010

Repositorios Digitales con Contenidos Orientados a las Necesidades de Instituciones Educativas de la Región

Antonio Castro Lechtaler^{1,2}, Fernanda Carmona¹, José Texier¹, Fernando Emmanuel Frati¹, Alberto Riba¹, Marisa Gagliardi¹, Emiliano Peressini¹, Germán Kurt Grin²

¹Departamento de Ciencias Básicas y Tecnológicas, Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina; ²CISTIC/FCE - Universidad de Buenos Aires, Córdoba 2122, CABA

antonio.castro@fce.uba.ar, {fbcarmona, jtexier, fefrati, ariba}@undec.edu.ar,
marisagagliardi76@gmail.com, emiliano.peressini@gmail.com, german.grin@gmail.com

RESUMEN

Las políticas públicas como los programas “Conectar Igualdad” y “Joaquín V. González”, han alcanzado una amplia cobertura en la entrega de computadoras portátiles y han contribuido en cambios educativos y sociales en la Argentina, pero su impacto no ha sido el esperado por diversas causas en las Instituciones Educativas (IE). Adicionalmente, se aprobó la Ley 26.899 de Repositorios Institucionales (RI) y la Ley 26.917 sobre el Sistema Nacional de Bibliotecas Escolares. Esta línea de investigación propone una solución a las diferentes necesidades escolares presentes en las IE de la región, relacionadas con problemas en el acceso a bibliografía básica y técnica, para contribuir, sin costo alguno, en la formación del alumno y en la capacitación del RRHH de las mismas. Los Repositorios Institucionales (RI) pueden brindar un espacio de contención a obras de referencia, contenidos multimediales, televisivos, libros, imágenes, videos, manuales áulicos y otros recursos educativos adecuados a estos niveles iniciales. La implementación de un RI vinculado al requerimiento de otros espacios académicos constituye una oportunidad de articulación entre todos los niveles educativos, mediante el aporte tecnológico que caracterizan el espacio universitario y respaldado por el marco normativo vigente. En este sentido, tanto el Centro de Investigación en Sistemas, Tecnologías de la

Información y las Telecomunicaciones de la Universidad de Buenos Aires, como el Departamento de Ciencias Básicas de la Universidad Nacional de Chilecito han buscado articular sus recursos y producciones científicas – tecnológicas, garantizando un conjunto de trabajos acordes a las líneas de investigación de ambas instituciones. Dentro de esas líneas de trabajo se encuentran el desarrollo de Repositorios Institucionales (RI), Objetos de Aprendizaje (OA) y el movimiento de Acceso Abierto (AA); tanto en lo que respecta a los aspectos técnicos vinculados a ellos, como a su utilización y uso como productos útiles en la enseñanza primaria y secundaria.

Palabras clave: repositorios institucionales, objetos de aprendizaje, material educativo digital, acceso abierto.

CONTEXTO

En el marco del Plan de Fortalecimiento de la Investigación Científica, el Desarrollo Tecnológico y la Innovación en las Universidades Nacionales (Ac. Pl. N° 676/08 y N° 687/09) y del Reglamento PDTS–CIN (Ac. Pl. N° 901/14), el Consejo Interuniversitario Nacional y el CONICET convocan, en el año 2014, a Instituciones Universitarias Públicas a la presentación de Proyectos de Desarrollo Tecnológico y Social (PDTS) para la adjudicación de subsidios.

La convocatoria exigía que dentro del equipo de investigación del proyecto, la dirección fuese realizada a través de un Grupo Responsable, que debía incluir a docentes-investigadores de al menos dos Universidades, preferentemente de distintos CPRES (Consejos Regionales de Planificación de la Educación Superior). Dado que existía una coincidencia en las líneas de Investigación de ambas Universidades, la UBA presentó un proyecto conjuntamente con la UNDEC, el cual dentro del total de Proyectos presentados recibió el número de orden 407.

Esta línea de I/D/I corresponde al diseño, desarrollo e implementación de proyectos que fortalecerán la investigación y las diferentes labores relacionadas con la gestión del conocimiento en cuanto a la visibilidad web y la preservación de la mayoría de la producción académica y científica tanto de la UNdeC como de la FCE de la UBA y otras Instituciones Educativas de nivel primario y secundario, fortaleciendo la articulación entre todos los niveles educativos. Esta línea de investigación se encuentra alineada con los núcleos estratégicos del Plan Nacional CyT 2020 ya que comprende el desarrollo de bases de datos dirigidas a la generación de sistemas con fines socio-productivos inclusivos, al mejoramiento de las comunicaciones entre las comunidades y a la creación de Centros de Innovación Inclusiva.

En lo que respecta a la UNdeC cabe mencionar otros proyectos relacionados a esta línea de investigación como el proyecto “Red para la creación y publicación de objetos virtuales de aprendizajes de calidad en Repositorios Institucionales” que corresponde a la convocatoria “Redes Internacionales 9”, promovida por la SPU, período de ejecución 2016-2017; un proyecto presentado en la convocatoria 2013-2014 del programa “Financiamiento para el Estímulo y Desarrollo de la Investigación Científica y Tecnológica” de la Secretaría de Ciencia y Tecnología en la

UNdeC cuyo título es “Fortalecimiento y visibilidad web de la producción científica de la UNdeC”.

Los proyectos del programa de Redes Internacionales permiten generar los espacios para discutir, exponer y articular, en diferentes ámbitos nacionales e internacionales, el trabajo y los resultados obtenidos, mientras que el proyecto correspondiente al programa de Financiamiento para el Estímulo y Desarrollo de la Investigación Científica y Tecnológica sirve de base para la búsqueda de trabajos científicos que puedan colaborar y responder a las necesidades de las IE.

INTRODUCCIÓN

La incorporación de las Tecnologías de la Información y Comunicación (TIC) en la Educación ha permitido extender los ambientes de enseñanza y aprendizaje, para así poder desarrollar ambientes virtuales colaborativos e interactivos, con el uso de diferentes materiales educativos digitales, que permitan el logro de los objetivos. Es por ello, que en el ámbito educativo se ha impuesto un nuevo concepto que busca la reutilización, permanencia, interoperabilidad, accesibilidad y compatibilidad de recursos digitales para el desarrollo de cursos y programas de formación en línea a través de la Web, los llamados Objetos de Aprendizaje (OA). Lo que distingue a un OA de un material educativo digital es la introducción de información autodescriptiva, expresada a través de los metadatos, conjunto de atributos o elementos que permiten describir al objeto, es decir son entidades generalmente entregadas a través de Internet y diseñadas con el fin de que sean utilizadas y reutilizadas en múltiples contextos educativos [5, 6].

En los últimos años, el desarrollo de Repositorios Institucionales (RI) de acceso abierto ha sido un tema prioritario en las políticas de educación, ciencia y técnica de muchos países, y en particular, en las universidades públicas de Argentina,

aprobándose la Ley 26899 de Repositorios Institucionales (RI), y en los otros niveles educativos las Bibliotecas Escolares con la aprobación de la Ley 26.917 sobre el Sistema Nacional de Bibliotecas Escolares y Unidades de Información Educativa.

Los RI tienen como propósito recopilar, catalogar, gestionar, acceder, difundir y preservar información, permitiendo entre otras cosas el acceso libre y gratuito a todos los recursos que los conforman [1], [2]. Estos están constituidos por un conjunto de archivos digitales en representación de productos científicos y académicos que pueden ser accedidos por los usuarios [1]. En este sentido, un RI puede integrar a los OA en los denominados Repositorios de Objetos de Aprendizaje (ROA), bibliotecas digitales especializadas, orientados a facilitar la búsqueda y recuperación de los OA de manera que puedan ser utilizados en diversos ambientes educativos [11].

La filosofía del Acceso Abierto tiene como fin asegurar el acceso libre y abierto a la producción científica y académica, es decir, garantiza el acceso a través de Internet sin que los derechos del copyright sean una barrera [3], [4]. Es así que los RI materializan el objetivo del OA porque la información que se deposita es una producción que tiene como propósito ser accesible sin restricciones, y preservada digitalmente.

El RI permitirá contar en la región con material de estudio generado por instituciones de nivel primario y secundario, contribuyendo a la articulación de todos los niveles. El compartir los trabajos, obras, documentos, libros, trabajos prácticos, entre otros, generados por los actores, permitirá mejorar la calidad de la educación que recibirán los estudiantes y aumentar la eficiencia de la investigación. Esa información podrá ser utilizada para explorar nuevas hipótesis, conocimientos previos o nuevos, validar métodos de es-

tudio, técnicas de análisis o estrategias de aprendizaje utilizadas en las distintas IE bajo el principio de acceso a la información libre y gratuita, es decir, sin restricciones legales, técnicas y de acceso.

A través de los Centros de Innovación Inclusiva creados, todos los integrantes de la comunidad podrán constituirse no sólo en consumidores de saberes, sino en creadores de conocimiento. Aunque la brecha digital se está reduciendo gracias a las políticas de inclusión nacionales y provinciales, aún existen sectores sociales con dificultades para acceder a la bibliografía básica y técnica necesaria para la formación. Estos sectores serán los principales beneficiarios del RI, ya que con el apoyo directo y trabajo conjunto de las IE también se tendrá un impacto indirecto en la comunidad de la que forman parte. Serán los mismos estudiantes que adquieran las competencias necesarias para superar el alfabetismo digital y son quienes contribuirán desde sus hogares, escuelas o cualquier otro lugar para replicar estos resultados, sobre todo fomentando el hábito de la lectura/escritura en situaciones donde el acceso físico al material bibliográfico suele ser escaso. Esto dará lugar a nuevos proyectos orientados a esos espacios de la comunidad, dirigidos precisamente a superar sus dificultades de inclusión digital.

LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Este trabajo incluye las siguientes líneas de investigación:

- Repositorios institucionales y bibliotecas digitales.
- Objetos de aprendizaje.
- Gestión de la información y el conocimiento.
- Sistemas de información web y bases de datos.
- Índices bibliométricos.
- Interoperabilidad.
- Preservación digital.

- Recuperación de la información.

OBJETIVOS

El desarrollo de esta línea de investigación está permitiendo cumplir con los siguientes objetivos:

- Comprender los problemas y las tendencias asociadas a la producción abierta del conocimiento, el uso y la difusión.
- Analizar la visibilidad web de las IE intervinientes.
- Analizar los recursos educativos existentes que cumplan con las necesidades educativas de las IE intervinientes.
- Fomentar la formación de habilidades y conocimientos relacionados con los RI y los OA.
- Entender y valorar el movimiento mundial de Acceso Abierto y señalar las vías de su materialización.
- Implementar Repositorios de prueba para visualizar los diferentes recursos (educativos, académicos y científicos) que se produzcan en las IE de la región.
- Depositar y preservar los OA generados por los diferentes actores de las IE.
- Definir e implementar políticas de recopilación, distribución y mantenimiento para el funcionamiento del RI.
- Establecer principios y normas para que permitan la implementación de los Centros de Innovación Inclusiva [7, 8].
- Contar en la región con el primer RI con material educativo digital generado por IE de nivel primario y secundario.

RESULTADOS

OBTENIDOS/ESPERADOS

- Relevamiento de los diferentes recursos educativos producidos y a producir por las IE, con especial énfasis en las escuelas rurales de montaña, e incluyendo además, a todas las instituciones educativas de la ciudad de Chilecito y su zona de influencia, siguiendo los li-

neamientos del proyecto PDTs, anteriormente nombrando.

- Curso a nivel de postgrado llamado “Gestión de la Información Académica y Científica”.
- Desarrollo de un proyecto de asignaturas sobre las Bibliotecas Digitales para presentarse en dos universidades [9, 10] y otras que lo puedan requerir.
- Implementación de un repositorio de prueba que está gestionando los recursos educativos y objetos de aprendizaje generados dentro del proyecto PDTs.
- Desarrollo, en las IE de nivel primario y secundario, de talleres sobre herramientas de ofimática, de producción colaborativa de documentos, de búsqueda y selección de información en la web.
- Desarrollo, en las IE de nivel primario y secundario, de talleres para la formación de habilidades y conocimientos relacionados con los RI y OA y producción de OA.
- Certificación de los cursos realizados.

FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por investigadores categorizados y otros en formación del CISTIC/FCE y de la Carrera de Licenciatura en Sistemas de Información de las Organizaciones Universidad de Buenos Aires por una parte; y por la otra, de investigadores doctorados y categorizados pertenecientes a de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas de la UNdeC, especializados bibliotecas digitales, repositorios institucionales, objetos de aprendizaje, desarrollo de software, cómputo paralelo y tecnología grid; participando también alumnos avanzados de grado de ambas Universidades.

Por lo expresado, se cuenta con los recursos humanos con las habilidades y la formación académica en las diversas áreas de

la propuesta, asegurando la concreción de esta línea de trabajo.

Los integrantes son docentes de las asignaturas las siguientes asignaturas:

En la Universidad de Buenos Aires: Teoría de los Lenguajes y Sistemas Operativos, Construcción de Aplicaciones Informáticas, y Sistemas de Datos.

En la Universidad Nacional de Chilecito: Programación I, Sistemas I, Arquitecturas Paralelas, Teoría de la Computación y Herramientas de Ingeniería de Software.

BIBLIOGRAFÍA

- [1]. J. Texier, “Los repositorios institucionales y las bibliotecas digitales: una somera revisión bibliográfica y su relación en la educación superior,” presented at the 11th Latin American and Caribbean Conference for Engineering and Technology - 2013, Cancun, México, 2013, p. 9.
- [2]. M. De Giusti, N. Oviedo, A. Lira, A. Sobrado, J. Martínez, and A. Pinto, “SEDICI – Desafíos y experiencias en la vida de un repositorio digital,” *RENATA*, vol. 1, no. 2, pp. 16–33, Aug. 2011.
- [3]. P. Suber, “Ensuring open access for publicly funded research,” *BMJ*, vol. 345, 2012.
- [4]. D. Torres-Salinas, N. Robinson-García, and A. Cabezas-Clavijo, “Compartir los datos de investigación en ciencia: introducción al data sharing,” *Profesional de la Información*, vol. 21, no. 2, pp. 173–184, 2012.
- [5]. D. Wiley, “Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy,” presented at the Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos, SPEDECE, 2001.
- [6]. E. Morales, F. García, A. Barron, A. Berlanga, and C. López, “Propuesta de Evaluación de Objetos de Aprendizaje,” presented at the Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos Reutilizables (SPDECE), Barcelona, 2005.
- [7]. La Innovación Inclusiva: <http://www.mincyt.gob.ar/noticias/una-jornada-para-pensar-sobre-la-innovacion-inclusiva-12550>.
- [8]. La Innovación al Servicio de la Inclusión Social: <http://www.mincyt.gob.ar/noticias/la-innovacion-al-servicio-de-la-inclusion-social-10731>
- [9]. Texier, J., Zambrano, J., & Carmona, F. B. (2016). Las Bibliotecas Digitales en el Currículum de las carreras de Ciencias de la Computación: una propuesta posible para Argentina y Venezuela. In XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016).
- [10]. Texier, J. (2016). Los repositorios institucionales y su importancia en las Universidades Venezolanas. *Revista Scitus*. Venezuela.
- [11]. Paulsson, F., Naeve, A. Establishing technical quality criteria for Learning Objects. Proceedings of To be published in the proceedings of eChallenges 2006, Barcelona, Spain.

Análisis y detección de patrones en un grafo conceptual construido a partir de respuestas escritas en forma textual a preguntas sobre un tema específico.

María Alejandra Paz Menvielle, Cynthia Lorena Corso, Karina Ligorria, Analía Guzmán, Martín Casatti

Departamento de Ingeniería en Sistemas de Información
Facultad Regional Córdoba – Universidad Tecnológica Nacional
Maestro Marcelo López esq. Cruz Roja Argentina – Córdoba
0351 – 4686385

pazmalejandra@gmail.com, cynthia.corso@gmail.com, karinaligorria@hotmail.com,
aguzman@sistemas.frc.utn.edu.ar, mcasatti@gmail.com,

Resumen

En este trabajo se presenta una línea de Investigación que busca descubrir patrones asociados a la evolución de una base de conocimiento representada en una base de datos orientada a grafos, la misma contiene respuestas de exámenes en formato de texto de redacción libre relacionadas a un dominio específico, utilizada para realizar el análisis de texto en respuestas a preguntas de exámenes en la cátedra de Paradigmas de Programación, con el objetivo de detectar el grado de acierto de las respuestas de los alumnos. Dentro de los patrones que se pretenden descubrir se encuentran aquellos asociados a las respuestas de los alumnos, a la forma de representación de las preguntas de los docentes, entre otras. Es por ello que el presente proyecto busca avanzar en la línea de investigación relacionada a la detección

de patrones a partir de grafos dirigidos, tanto en sus aspectos teóricos como prácticos y en sus aplicaciones.

Palabras clave: grafos – patrones – rutas – comunidades.

Contexto

El presente trabajo forma parte del proyecto de investigación y desarrollo que ha sido homologado por la Secretaría de Investigación, Desarrollo y Posgrado de la Universidad Tecnológica Nacional, reconocido con el código PIDEIUTNCO0004812, el mismo forma parte del Centro de Investigaciones, Desarrollo y Transferencia de Sistemas de Información – CIDS.

Para su desarrollo se utilizará como caso testigo a la cátedra de Paradigmas

de Programación, perteneciente a la carrera Ingeniería en Sistemas de Información, dictada en la Facultad Regional Córdoba, de la Universidad Tecnológica Nacional. Se respetarán los contenidos mínimos fijados para esta asignatura tal cual figuran en la ordenanza 1150 de la carrera, los cuales pertenecen al bloque de tecnologías básicas dentro del área programación, que están principalmente referidos a los paradigmas lógicos, funcional y de orientación a objetos. Además se cumplirá con los descriptores y criterios de intensidad de formación práctica de la Resolución Ministerial 786/09, los que se encuentran definidos en el área de tecnologías básicas, sub-área programación que incluyen a los paradigmas y lenguajes de programación.

El trabajo que aquí se presenta es la continuación de los trabajos realizados durante el desarrollo del PID EIUTNCO0003592 “Metodología para determinar la exactitud de una respuesta, escrita en forma textual, a un interrogatorio sobre un tema específico”. Durante el transcurso de dicho proyecto se generó una base de datos de grafo cuyo objetivo principal es la registración y almacenamiento de todos los conceptos contenidos en la curricula de la materia Paradigmas de Programación. En el presente proyecto se propone complementar la funcionalidad del proyecto anterior, mediante la búsqueda, el análisis y la propuesta de patrones topológicos frecuentes en un grafo conceptual construido para determinar la exactitud de las respuestas, escritas en forma textual sobre un tema específico.

Introducción

Un patrón es una entidad a la que se le puede dar un nombre y que está representada por un conjunto de propiedades medidas y las relaciones entre ellas (*vector de características*) [1].

En el campo del Reconocimiento de Patrones un enfoque que está ganando popularidad es el de la aplicación de grafos como herramienta para la representación de entidades con estructuras complejas [2].

En una representación de este tipo los vértices y sus atributos representan objetos (o partes de ellos) mientras que los arcos representan relaciones entre estos objetos. Este enfoque explota la generalidad inherente de las representaciones basadas en grafos y gracias a las mejoras en la capacidad de procesamiento de las computadoras, estas representaciones estructurales y los algoritmos que sobre ellas se aplican se han vuelto más eficientes en su aplicación [2].

Caracterización

A la hora de determinar patrones dentro de un grafo dirigido existen un conjunto de medidas que caracterizan el grafo y que determinan qué tipos de técnicas se pueden utilizar de manera más o menos eficiente para obtener información estructural del mismo [3]. Estas medidas se denominan “características” o “métricas” y son inherentes al grafo en su conjunto, son dinámicas y cambian a medida que nuevos nodos y arcos se van incorporando a la estructura. Algunas de las

características más importantes de los grafos son:

Tamaño: El tamaño de un grafo se determina por la cantidad de nodos que lo componen.

Grado de un vértice: Es la cantidad de arcos que convergen en el mismo. Esta característica es particular de cada vértice pero se utilizan algunas medidas generales, como el grado máximo, mínimo o promedio, para caracterizar el grafo de manera general.

Densidad: La densidad es la relación existente entre la cantidad total de arcos del grafo con respecto a los nodos que lo conforman. Un grafo denso, en este contexto, es un grafo con una gran cantidad de interconexiones entre nodos.

Isomorfismo en grafos

Consideramos un grafo G definido como $G = (V, E, L)$ siendo V un conjunto finito de vértices, E un conjunto finito de arcos entre los mismos y L un conjunto finito de etiquetas que se pueden aplicar tanto a vértices como a arcos.

Un grafo $G' = (V', E', L')$ es isomórfico de G si: Existe un mapeo biyectivo entre los vértices V y V' . Si existe un arco E entre dos vértices de G existe también un arco E' entre los dos vértices correspondientes en G' . Las etiquetas utilizadas en G son preservadas al realizar el mapeo entre G y G' [4].

Especificación de patrones

Una forma simple de detección de patrones en grafos es el problema de encontrar un subgrafo (el “resultado”) de un grafo de entrada dado (el “objetivo”) tal que ese subgrafo sea

isomórfico de otro grafo de entrada (el “patrón”) [5].

Un enfoque más general busca encontrar todos los subgrafos (“resultados”) dentro del grafo objetivo y no solamente uno de ellos.

Análisis de patrones comunes

Mientras que la minería de datos se enfoca principalmente en los valores de los datos que se están buscando, en los esquemas semi-estructurados y en los grafos, el enfoque se encuentra en etiquetas frecuentes y topologías comunes [6]. En estos la estructura de los datos es tan importante como su contenido.

Originalmente se plantearon soluciones para el hallazgo de estructuras representadas por una ruta simple (single-path) y para estructuras de tipo árbol, pero actualmente muchas de las estructuras que se encuentran en la Web, así como en redes sociales o comunidades online tienen la forma de grafos más complejos, tanto cíclicos como acíclicos.

Es por eso que ésta disciplina ha experimentado un resurgimiento debido a que el descubrimiento de los patrones subyacentes posibilita un mejor diseño de las bases de datos que gestionan estas estructuras y un mejor indexamiento en la aplicación de algoritmos que tienen en cuenta las preferencias de los usuarios (recomendaciones online, compras, sugerencias de grupos afines, etc.), en las predicciones de comportamiento, entre otras.

Si bien surgen y se mejoran algoritmos de uso específico para detección de patrones conocidos, un problema inherente a éstas técnicas, radica en el

planteo de algoritmos de uso general (que presenten un rendimiento razonable) y en el descubrimiento de patrones, hasta el momento desconocidos mediante técnicas generales [9].

Líneas de Investigación, Desarrollo e Innovación

- Estudio de los diferentes patrones topológicos de grafos que puedan ser relevantes en la búsqueda de información en el dominio elegido, analizando si dichos patrones tienen comportamientos recurrentes o subyacentes.
- Estudio de algoritmos que permitan detectar patrones conocidos en la teoría de grafos como son las “comunidades”, “pares”, “rutas principales” y otros patrones comunes por medio del análisis de las métricas de la base de conocimientos.
- Detección de patrones que, aún no siendo comunes en otras áreas de la teoría de grafos, si lo son recurrentes en el dominio bajo estudio.
- Aplicación de los patrones encontrados sobre las respuestas elaboradas por los alumnos en un examen posibilitará descubrir algunas características importantes que se relacionan con el aprendizaje [7].
- Estudio de herramientas de visualización y análisis de grafos (como Gephi o GUESS, entre otras), para realizar los análisis preliminares y la determinación de

los parámetros y métricas de la base de datos [8].

- Automatización de algunos de estos análisis para incluirlos en una herramienta ad-hoc.

Resultados y Objetivos

El objetivo del presente estudio es analizar, detectar y evaluar patrones topológicos frecuentes en un grafo conceptual construido para determinar la exactitud de las respuestas, escritas en forma textual sobre un tema específico, utilizando una base de conocimientos diseñada como un grafo dirigido. Para ello hemos identificado los siguientes objetivos particulares:

1. Explorar patrones topológicos de grafos que contengan información relevante para la identificación de estructuras dentro de la base de conocimientos de la materia Paradigmas de Programación.
2. Analizar la existencia de patrones recurrentes o subyacentes en los grafos generados a partir de las respuestas base de los docentes y los obtenidos de las respuestas dadas por los alumnos.
3. Proponer algoritmos que permitan detectar patrones conocidos en la teoría de grafos como son las "comunidades", "pares", "rutas principales" y otros patrones comunes, por medios del análisis de las métricas sobre la base de conocimiento.
4. Identificar características como exactitud, coherencia y consistencia, entre otras, de las respuestas escritas en forma textual, en la base de conocimiento diseñada como un grafo dirigido.

Formación de Recursos Humanos

Dentro del desarrollo de este proyecto de investigación se está desarrollando el trabajo de Tesis de Maestría de dos integrantes docentes del presente proyecto. Se incorporan al equipo de trabajo docentes-investigadores de la carrera Ingeniería en Sistemas de Información como investigadores de apoyo con la finalidad de que inicie su formación en investigación científica y tecnológica, se incorpora un becario graduado BINID y becarios alumnos, quienes colaborarán en la recolección, manipulación y desarrollo de este marco metodológico. En el marco del proyecto los estudiantes tendrán la posibilidad de hacer la Práctica Supervisada de quinto año. Los avances, propuestas y herramientas construidas, estarán disponibles para su transferencia y aplicación en el Centro de Investigaciones, Desarrollo y Transferencia de Sistemas de Información - CIDS. Del mismo modo la detección de patrones sobre el dominio de conocimiento de la materia Paradigmas de Programación continuará beneficiando a los integrantes de la cátedra y a los estudiantes.

Referencias

[1] Watanabe, Satoshi. Pattern recognition: human and mechanical. John Wiley & Sons, Inc., 1985.

[2] "Graph for pattern recognition" (Doctoral dissertation, Ph. D. dissertation, Universite Francois-Rabelais). Raveaux, R., & Abu, Z. (2013).

[3] Joyner, David, Minh Van Nguyen, and David Phillips. "Algorithmic graph theory and sage." *Version 0.8-r1991* (2013)

[4] VAN STEEN, Maarten. An Introduction to Graph Theory and Complex Networks. *Copyrighted material*, 2010.

[5] "An algorithm for subgraph isomorphism". J. R. Ullmann. Journal ACM, 23(1):31-42, 1976.

[6] "Computing frequent graph patterns from semistructured data. En Data Mining". VANETIK, Natalia; GUDES, Ehud; SHIMONY, Solomon Eyal. 2002. ICDM 2003. Proceedings. 2002 IEEE International Conference on. IEEE, 2002. p. 458-465.

[7] "Introduction to information retrieval". C. D. Manning, Prabhakar Raghavan, and Hinrich Schütze. Cambridge University Press. Julio 2008. También disponible on line: www.safaribooksonline.com

[8] "Classification of graph metrics". Hernández, Javier Martín, and Piet Van Mieghem. Delft University of Technology, Tech. Rep (2011).

[9] FOGGIA, Pasquale; SANSONE, Carlo; VENTO, Mario. A performance comparison of five algorithms for graph isomorphism. En *Proceedings of the 3rd IAPR TC-15 Workshop on Graph-based Representations in Pattern Recognition*. 2001. p. 188-199.

Caracterización preliminar de los aspirantes a ingresar a la Licenciatura en Ciencias de la Computación

Carlos García Garino^{1,2}, Marisa Haderne¹, Lucía Cortés¹, Carlos Catania^{1,2}, Elina Pacini^{1,2,3}

¹ Facultad de Ingeniería. Universidad Nacional de Cuyo.

² ITIC, Universidad Nacional de Cuyo.

³ CONICET, Carrera del Investigador Científico Técnico

cgarcia@itu.uncu.edu.ar, hadernemarisa@yahoo.com.ar; luciacortes5519@gmail.com, ccatania@itu.uncu.edu.ar, epacini@uncu.edu.ar

RESUMEN

La deserción de los estudiantes universitarios en los primeros años de la carrera es una preocupación presente en todas las instituciones. Por este motivo, es importante analizar el problema en una etapa previa al cursado de la carrera, cuando el estudiante está transitando su ingreso a la institución seleccionada. Los resultados obtenidos pueden luego ser contrastados con el rendimiento de los estudiantes en los primeros años de la Universidad.

El presente trabajo apunta a caracterizar el comportamiento de los aspirantes a ingresar en la Licenciatura en Ciencias de la Computación, mediante técnicas de minería de datos. Para ello, se dispone de datos censales y rendimiento durante el cursado del ingreso de las cohortes 2017 y 2018.

Para explicar el problema se han propuesto diversas variables: situación socioeconómica, condición cultural, institución en la cual cursó estudios medios, rendimiento en los exámenes de ingreso, motivación personal, etc. Todas ellas comprendidas dentro de las características y comportamientos propios del aspirante universitario.

En este trabajo se presentan los resultados obtenidos y se discuten los resultados esperados.

Palabras clave: 1) Deserción estudiantil 2) Minería de Datos Educativa 3) Rendimiento de los estudiantes

CONTEXTO

En la Figura 1 se muestran las líneas de investigación que se llevan a cabo en el Instituto Universitario para las Tecnologías de la Información y las Comunicaciones (ITIC), y que se comparten con la Carrera de Licenciatura en Ciencias de la Computación, que se dicta en la Facultad de Ingeniería de la Universidad Nacional de Cuyo (UNCuyo).

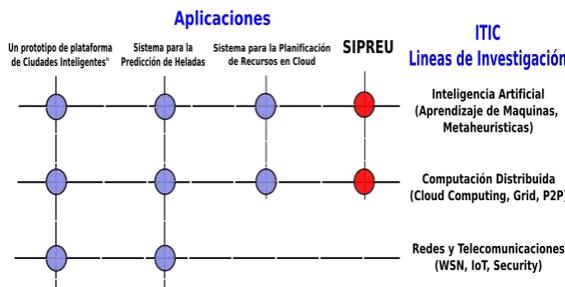


Figura 1: líneas de investigación del ITIC y carrera de Licenciatura en Ciencias de la Computación

En este trabajo se presentan los avances del proyecto “SIPREU: Un sistema inteligente para caracterizar y predecir el rendimiento de estudiantes universitarios” [7], cuyos lineamientos se comunicaron en el congreso WICC 2017 [9] y en la conferencia ADNTIIC [8]. En el proyecto PIN “La Minería de Datos como herramienta para estudiar el rendimiento de estudiantes de la Licenciatura

en Ciencias de la Computación” [2], también se estudian algunos aspectos de interés del problema planteado.

1. INTRODUCCIÓN

El problema planteado en el presente trabajo es relevante y de especial interés para la Facultad de Ingeniería de la Universidad Nacional de Cuyo. Es fundamental para la Facultad conocer al aspirante de la Licenciatura en Ciencias de la Computación, su problemática académica, trayectos educativos anteriores y otros componentes sociales que pueden afectar su rendimiento al momento del ingreso.

La carrera es reciente, su primera cohorte comenzó en el 2017 y acaba de iniciarse el cursado de la segunda cohorte. Para las autoridades es muy importante analizar y conocer las características de sus ingresantes desde el inicio de la carrera.

Por otra parte, el problema posee mucha relevancia local ya que se enmarca dentro del Objetivo 1 del Plan Estratégico 2021 de la UNCuyo [11].

En un trabajo anterior [7] se discutieron con detalle trabajos relacionados con la presente propuesta. A manera de síntesis puede mencionarse el trabajo de Garbanzo Vargas [6] que propone agrupar a los diferentes factores que caracterizan el comportamiento de un estudiante en tres grupos: a) Determinantes personales; b) Determinantes sociales y c) Determinantes institucionales.

Dado que muchos de los factores no se pueden medir con facilidad en el momento del acceso del estudiante a la Universidad, se tendrán en cuenta los factores propios de los determinantes sociales, ya que los mismos parecen, según estudios realizados, brindar un buen punto de partida.

En cualquier caso, los factores (atributos) para caracterizar la población inicial surgirán de los estudios mencionados y de una discusión interdisciplinar entre los integrantes del equipo de trabajo, entre quienes se encuentran especialistas en informática, postgraduados en educación, profesores universitarios, etc.

También se ha registrado el rendimiento de los aspirantes (cohortes 2017 y 2018) a

ingresar a la carrera de Licenciatura en Ciencias de la Computación. La información recogida es una muestra reducida, pero suficiente para brindar algunos indicadores útiles para avanzar en el proyecto, como se discute en la sección 3 de este trabajo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Además de la relevancia general y local del problema corresponde analizar la relevancia disciplinar del tema. La Minería de Datos Educativa es una línea que ha cobrado mucha relevancia mundial en los últimos años, como se observa en el portal dedicado al tema que lleva adelante la International Educational Data Mining Society [3]. En dicho sitio se expresa textualmente: “La Minería de Datos en educación es una disciplina emergente, orientada al desarrollo de métodos para explorar el creciente volumen de datos provenientes de los sistemas de educación, y utilizar esos métodos para comprender mejor a los estudiantes y sus entornos de aprendizaje”.

La citada sociedad edita un journal específico: Journal of Educational Data Mining y lleva adelante las conferencias internacionales acerca del tema, cuya próxima edición tendrá lugar en Buffalo, Nueva York en 2018.

3. RESULTADOS OBTENIDOS / ESPERADOS

Datos censales y de rendimiento

Se dispone de datos censales y de rendimiento de los aspirantes a ingresar en la carrera de Licenciatura en Ciencias de la Computación, correspondientes a las cohortes 2017 y 2018. Como se observa en la Figura 2, ingresaron 32 alumnos a la cohorte 2017 de la carrera.

La cohorte 2017 presentó la particularidad que 16 estudiantes (14%) ingresaron a través del otorgamiento de equivalencias, mientras que los otros 15 (13%) de los aspirantes aprobaron el curso de ingreso. El 73% restante se distribuye entre quienes

desaprobaron el ingreso (47%) y quienes no registraron actividad (26%).

Los resultados obtenidos reflejan una tendencia habitual en las carreras nuevas que incorporan muchos estudiantes con experiencias universitarias previas.

Para la cohorte 2018, como se muestra en la Figura 3, de un total de 110 aspirantes, aprobaron el ingreso 32 alumnos (29%).

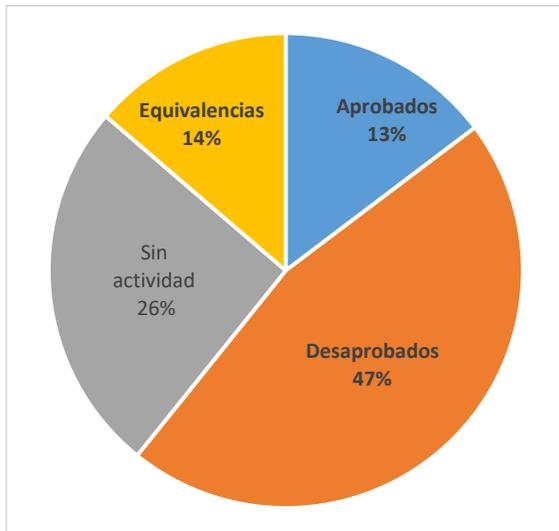


Figura 2: Cohorte 2017. Rendimiento de los aspirantes a ingresar en la Carrera de Licenciatura en Ciencias de la Computación.

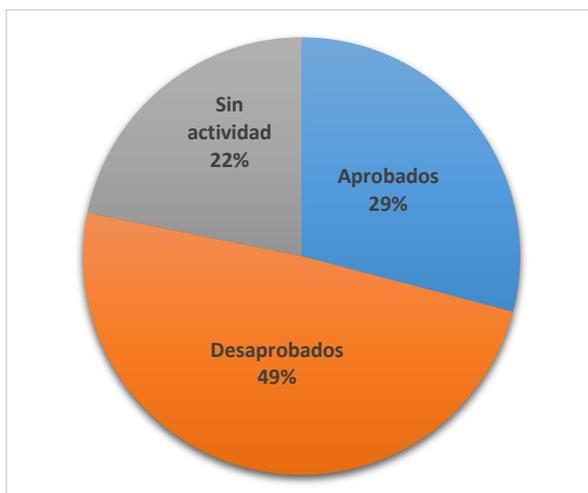


Figura 3: Cohorte 2018. Rendimiento de los aspirantes a ingresar en la Carrera de Licenciatura en Ciencias de la Computación.

Determinación de la muestra

Para determinar la muestra se integraron los archivos de datos personales y de calificaciones obtenidas por los aspirantes a ingresar en 2018. A partir del archivo unificado se hizo un análisis exhaustivo de los datos para determinar los atributos que serán tenidos en cuenta para trabajar en el modelo descriptivo/predictivo.

Del análisis que se llevó a cabo surge:

1. Se descartan los casos que no registran actividad, es decir, aspirantes inscriptos que no se presentaron a ninguna instancia de evaluación.
2. De las calificaciones obtenidas se observa que de los alumnos que aprobaron Matemática también aprobaron Resolución de Problemas, y no viceversa. Por este motivo se ha optado por tener en cuenta como atributo sólo el promedio de calificaciones obtenidas en Matemática.
3. También se descartan atributos repetitivos como País y Provincia de procedencia y Domicilio por protección de datos personales.
4. Existen instancias de las que no se disponen algunos atributos como Departamento y Localidad aunque en algunos casos pudieron completarse tomando en cuenta la característica del número telefónico.
5. En conclusión, la vista minable se integra de 116 instancias con 7 atributos: Sexo, Colegio Secundario, Tipo de escuela secundaria (Gestión privada o gestión pública: UNCuyo, Cens, Técnicas, Otros estatales) Departamento, Localidad, Promedio, Ingresó (SI/NO).
6. De las 116 instancias, en 47 se verifica "Si ingresó" y 69 "No ingresó", por lo tanto es posible a partir de esta muestra conseguir un modelo apropiado.

Resultados obtenidos

A partir de la muestra previamente descrita se realizó un análisis tipo exploratorio de los datos. Para ello, se utilizaron algunas

herramientas visuales propias de la estadística como barplots y boxplots. Este análisis se llevó a cabo con el software R [12].

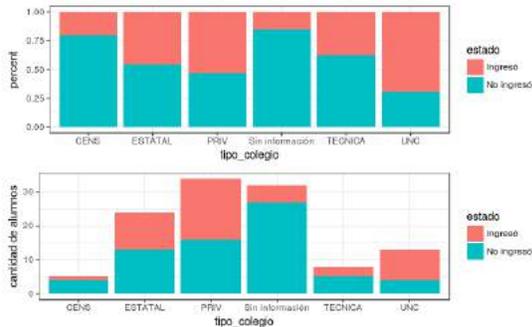


Figura 4: Distribución de la cantidad de aspirantes por tipo de colegio

En la figura 4, a priori no se observa diferencia entre los aspirantes provenientes de escuelas privadas y públicas. Si se observa un alto índice de aprobación en los aspirantes que provienen de colegios secundarios de la Universidad Nacional de Cuyo, indicados en los gráficos como UNC.

Posteriormente, se realizó el modelo de agrupación en clústeres de la muestra con la aplicación de Microsoft Clustering [10]. Este asistente divide los datos en un conjunto de entrenamiento (80%) y en un conjunto de pruebas (20%). Luego, se usa el conjunto de entrenamiento para entrenar los modelos y el conjunto de prueba para probar la precisión.

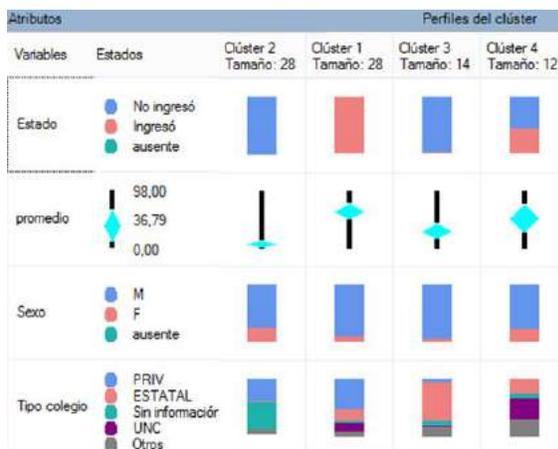


Figura 5. Modelo de agrupación en clústeres

Se puede observar en la Figura 5 que el cluster 1 posee la mayor cantidad de ingresantes, 95%, siendo el 54% de escuelas privadas, 20% estatales y 16% de la UNCuyo

Resultados esperados

Al final del proyecto se espera contar con mayor información acerca de los estudiantes de las cohortes 2017 y 2018. Para ello se espera enriquecer la información censal disponible mediante entrevistas personales a partir de un cuestionario actualmente en elaboración.

Además, se contrastará el rendimiento de los ingresantes con el correspondiente desempeño en primer año, y se podrán comparar los resultados de este proyecto con otros obtenidos mediante técnicas estadísticas [3,4]. Por otra parte, se ha comenzado a trabajar en la articulación horizontal entre dos asignaturas de primer año, concretamente entre Geometría Analítica e Introducción a la Programación y a contrastar el rendimiento de los estudiantes en dichas asignaturas. Algunos resultados preliminares [13] se comunicaron durante el congreso ENIDI. Finalmente se espera obtener información propia de los alumnos a partir de los trabajos de Brottier [1] que ha estudiado a las nuevas generaciones de estudiantes.

4. FORMACIÓN DE RECURSOS HUMANOS

La modalidad del proyecto SIPREU [5] financiado por la UNCuyo tiene, entre otros objetivos, la formación de Recursos Humanos. Una de las integrantes del proyecto está próxima a presentar su proyecto de tesis doctoral bajo la dirección de los investigadores formados del mismo. Otra integrante del equipo de trabajo actualmente lleva a cabo un proyecto de investigador novel PIN [2], financiado por la Facultad de Ingeniería de la Universidad Nacional de Cuyo, facilitando de esta forma su inicio en actividades de I+D.

Además, distintos integrantes del equipo de trabajo del presente proyecto han trabajado en los últimos años en diversas áreas relacionadas con la caracterización de los alumnos en la Universidad Nacional de Cuyo. Se contará con su experiencia para alcanzar los objetivos propuestos en el proyecto

SIPREU, así como incrementar las publicaciones de sus integrantes.

5. BIBLIOGRAFIA

- [1] L. Brottier: Proyecto 06/B005 Expectativas laborales de las nuevas generaciones. Proyecto bienal, Universidad Nacional de Cuyo, 2013-2015.
- [2] L. Cortés: La Minería de Datos como herramienta para estudiar el rendimiento de estudiantes de la Licenciatura en Ciencias de la Computación. Proyecto de Investigadores Noveles. Facultad de Ingeniería, Universidad Nacional de Cuyo, convocatoria 2017.
- [3] Educational Data Mining Web Site: <http://www.educationaldatamining.org/>
- [4] D. Fernández: Proyecto 06/B188 Una mirada introspectiva de los indicadores académicos que influyen tanto en el desempeño y el desgranamiento de los estudiantes como en la duración real de la carrera. Proyecto bienal, Universidad Nacional de Cuyo, 2009-2011
- [5] D. Fernández: Proyecto 06B/246 Detección, identificación y análisis de los factores fundamentales, externos e internos, que influyen y definen el ingreso de los aspirantes a Ingeniería y Arquitectura de la Universidad Nacional de Cuyo. Proyecto bienal, Universidad Nacional de Cuyo, 2011-13
- [6] G. M. Garbanzo Vargas. Factores asociados al rendimiento académico en estudiantes universitarios, una reflexión desde la calidad de la educación superior pública: pp. 43-63, vol. 31, núm. 1, 2007, Educación. Universidad de Costa Rica. San Pedro, Montes de Oca. Disponible en: <http://www.redalyc.org/articulo.oa?id=44031103>
- [7] C. García Garino 06/B349 SIPREU: Un sistema inteligente para caracterizar y predecir el rendimiento de estudiantes universitarios. Proyecto Bienal Tipo II 2016 – 2018.
- [8] C. García Garino, M. Haderne, C. Catania, E. Pacini. SIPREU: An Intelligent System for characterization and Prediction of University Students performance. ADNTIIC 2016: Advances in New Technologies, Interactive Interfaces and Communicability. Córdoba.
- [9] M. Haderne, L. Cortés, C. García Garino, C. Catania y Elina Pacini: Caracterización del aspirante a ingresar a la Licenciatura en Ciencias de la Computación, pp. 231-235. Actas del XIX Workshop de investigadores en Ciencias de la Computación, WICC 2017. Buenos Aires, 27 y 28 de Abril de 2017. ITBA y Red UNCI. ISBN 978-987-42-5143-5.
- [10] Complementos de minería de datos de Microsoft SQL Server 2012 SP1 para Microsoft Office
- [11] Universidad Nacional de Cuyo: Plan estratégico 2021. Disponible en: <http://www.uncu.edu.ar/planificacion/upload/plan-estrategico-1.pdf>. 2014
- [12] R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

URL <https://www.R-project.org/>.

- [13] S. Raichman y E. Pacini. S. Raichman and E. Pacini. Intervención educativa de articulación entre las asignaturas Introducción a la Programación y Geometría Analítica. Noveno Encuentro de Investigadores y Docentes de Ingeniería EnIDI 2017, Ciudad de Mendoza, Argentina. En prensa

Implementación de un EVEA institucional para enriquecer la enseñanza de pregrado, grado y posgrado de la UNTDF

Depetris Beatriz, Zangara Alejandra, Feierherd Guillermo, Pendenti Horacio, Rodríguez Juan Manuel, Romano Lucas, Blanco Claudio, María Belén Aguilera, Rojas Sergio.

Instituto de Desarrollo Económico e Innovación
Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur
Hipolito Irigoyen 880 - Ushuaia - Tierra del Fuego
{bdepetris, gfeierherd, hpendenti, jrodriguez, lromano, cblanco, mbaguilera}@untdf.edu.ar
alejandra.zangara@gmail.com, sergiorc.ush@gmail.com

RESUMEN

Los Sistemas de Administración del Aprendizaje (LMS, por sus siglas en inglés) o EVEAs (Entornos Virtuales de Enseñanza y de Aprendizaje), utilizados en sus inicios para resolver los problemas que planteaba la Educación a Distancia, fueron rápidamente adoptados para complementar la Educación Presencial, dando origen a diferentes modalidades que extienden las posibilidades del aula presencial, incorporando opciones de enseñanza mediadas por las herramientas ofrecidas en estos entornos y modificando las modalidades educativas resultantes en el gradiente presencialidad-distancia.

En esta evolución han ido surgiendo un número significativo de EVEAs, que han logrado distintos niveles de trascendencia.

Las instituciones educativas de diferentes niveles se han ido apropiando de este tipo de herramientas, con grados cada vez mayores de convergencia áulica e institucional. Se han generado interfases entre EVEAs y sistemas de gestión institucional que permiten, además de gestionar los contenidos y la comunicación de cursos, gestionar planes curriculares completos y grandes cantidades de estudiantes. Dadas las características particulares de la UNTDF,

el uso de estas tecnologías resulta imprescindible. La selección de un EVEA institucional constituye el primer paso para avanzar ordenadamente en ese sentido.

El proyecto propone analizar un subconjunto de los EVEAs disponibles, en particular los de código abierto, considerando aspectos pedagógicos y tecnológicos, con el objetivo de recomendar el que mejor se adecue al modelo pedagógico planteado por la UNTDF [1]. Además, a partir de un relevamiento de los conocimientos del cuerpo docente sobre el tema, se contempla desarrollar un plan de capacitación y realizar un conjunto de experiencias piloto en distintas cátedras con asistencia del equipo de proyecto.

Palabras clave: EVEAs, evaluación de EVEAs, criterios de selección de EVEAs, mediación de la enseñanza, Chamilo

CONTEXTO

El presente proyecto forma parte del “Área de Investigación 8 - Desarrollo Informático”, del Instituto de Desarrollo Económico e Innovación de la Universidad Nacional de Tierra del Fuego (UNTDF).

El proyecto se presentó a la

convocatoria realizada por la UNTDF en septiembre de 2016 y resultó aprobado luego de haber sido sometido a evaluación externa. (Resolución Rectoral 060/2017 del 10/04/2017).

El proyecto se desarrolla formalmente hasta el 28/02/2019.

INTRODUCCIÓN

Un Entorno Virtual de Enseñanza y de Aprendizaje (EVEA) es una aplicación informática diseñada para facilitar la comunicación pedagógica entre los actores del proceso educativo [2]. Combina una variedad de herramientas con la finalidad de dar soporte a profesores y estudiantes y, al mismo tiempo, optimizar las distintas fases del proceso de enseñanza y de aprendizaje. Fueron diseñados con el propósito de facilitar la comunicación pedagógica entre los participantes de un proceso educativo, sea éste completamente a distancia, presencial, o de naturaleza mixta, es decir, un proceso que combina ambas modalidades en diversas proporciones (blended). Funcionan utilizando redes telemáticas como soporte, principalmente Internet. [3]

Si bien en un primer momento los EVEAs se utilizaron casi con exclusividad en la educación a distancia, hace mucho tiempo que se han incorporado muy fuertemente a la educación presencial, generando así las modalidades conocidas como aula extendida, aula invertida y, con un mayor grado de mediación, propuestas de “blended learning”. El propósito de un EVEA es mediar una propuesta educativa, organizar los contenidos en diferentes materiales de estudio, facilitar la comunicación entre los participantes de un proceso educativo (fundamentalmente entre docentes y alumnos, alumnos entre sí y docentes entre sí), realizar un

constante seguimiento de la realización de actividades intermedias y, si se decide, de la evaluación final.

Con el tiempo, la primera generación de entornos, cuyo objetivo principal (y casi único) era permitir la distribución de materiales y la evaluación mediante pruebas automáticas, ha dado paso a una segunda generación, inspirada en los nuevos conocimientos sobre cómo se aprende usando recursos en línea.

Por medio de estos entornos se ha logrado aprovechar las características de accesibilidad y cooperación entre los usuarios de la red, orientando fundamentalmente el proceso educativo hacia nuevas maneras de aprender, apoyadas en el trabajo colaborativo, acompañadas de una serie de recursos, más allá de la mera presencia del docente. [4]

En la actualidad existe un amplio abanico de entornos virtuales orientados a la educación superior. Los hay con características de software propietario, de software libre y en la nube (aunque en este último caso no siempre son considerados como plataformas LMS o EVEAs propiamente dichas, pues su mayor utilidad es la de permitir el apoyo a la clase presencial así como el desarrollo de MOOC (Cursos online abiertos y masivos). [5]

Las universidades y los docentes pueden elegir sus EVEAs para impartir un modelo de enseñanza E-Learning. Desde la aparición de la llamada Web 2.0 todos estos EVEAs han evolucionado, incluyendo nuevas herramientas colaborativas, como blogs, foros, wikis, etc. [6]

Cada uno de ellos cuenta con sus propias herramientas y funcionalidades. No es posible encontrarlas a todas en una única plataforma y, por esta razón, es importante evaluar cuáles resultan prioritarias teniendo en cuenta el modelo

educativo que se quiere implementar.

Si se consideran los desarrollos propios de distintas universidades, el número de EVEAs es numeroso. En algunos casos se habla de más de ciento treinta. Si nos limitamos a los que han logrado cierta preponderancia en el ámbito internacional encontramos al menos 40 de ellos. [7][8]

Todos ofrecen un conjunto común de prestaciones y otras que los diferencian. También evolucionan de distinta manera.

Teniendo en cuenta los distintos módulos, componentes, recursos o herramientas incorporadas, los entornos privilegian distintas funcionalidades. Actualmente es posible observar tres tendencias bien definidas en sus prestaciones: [9] citado en [3]

- Entornos centrados en la creación, gestión y distribución de contenidos que además incorporan algunas herramientas de comunicación.
- Entornos centrados en la comunicación y las actividades de enseñanza / aprendizaje que incluyen además herramientas para gestionar materiales.
- Entornos de trabajo en grupo para comunidades académicas que agregan algunas funcionalidades utilizables en la enseñanza.

Es evidente que detrás de estas características subyacen distintas concepciones del proceso pedagógico. Por lo tanto, entre las actividades a desarrollar se ha planteado indagar sobre las prácticas de enseñanza que realizan los docentes de la UNTDF, en especial las referidas a la inclusión de TICs. Esta indagación revelará el modelo de enseñanza en el que se basará la capacitación docente y la implementación del entorno seleccionado.

Por lo tanto un buen entorno debe ser lo suficientemente flexible como para

no imponer un modelo pedagógico particular, sino posibilitar la aplicación de la mayoría de ellos, a fin de que los docentes y los estudiantes lo experimenten como un aumento de posibilidades y no como una limitación, un elemento de distorsión o un freno a su creatividad y a su capacidad de innovación. [9] citado en [3].

Resulta claro entonces que no debe restringirse la selección solamente a los aspectos técnicos. Al respecto señala Clarenc: *“Es necesario contemplar a los LMS desde la perspectiva del aprendizaje y no desde la tecnológica, ya que los aspectos técnicos deben ser considerados recursos a ser utilizados en el proceso de implementación.”* [10]

Teniendo en cuenta esta diversidad la decisión de elegir un EVEA para uso institucional no es sencilla. Además de las cuestiones pedagógicas deben considerarse las económicas, pues, más allá de la gratuidad inicial de los desarrollados bajo licencias de fuente abierta, la inversión es importante y, en consecuencia, debe estar precedida por un proceso de evaluación destinado a valorar la calidad del EVEA y a determinar objetivamente las potencialidades y la proyección futura de cada uno.

A fin de formular un plan de implementación que permita el éxito del proyecto de introducción del EVEA elegido, debe tenerse en cuenta el estado actual de conocimientos del cuerpo académico. Esto permitirá desarrollar un plan de actualización de esos conocimientos así como los recursos necesarios para acompañar a las experiencias piloto.

1. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- Entornos virtuales de enseñanza y aprendizaje. Proyectos que los incorporan, metodologías, desarrollos, evaluación de su calidad y experiencias.
- Formación de recursos humanos en el uso de TICs.
- Trabajo colaborativo mediado por TICs. Conceptualización, análisis y desarrollo de software y metodologías.
- Dispositivos móviles como soporte para el trabajo colaborativo y para prácticas de aula extendida y blended- learning.

2. RESULTADOS OBTENIDOS / ESPERADOS

El proyecto se inició el 1º de marzo de 2017.

La 1ra. actividad del proyecto consistió en el relevamiento de las funcionalidades de los entornos a partir de la bibliografía existente y la experimentación con aquellos que estaban disponibles y que fueron considerados de interés para el proyecto.

Considerando las características citadas anteriormente, se investigaron un conjunto de plataformas de uso gratuito y de código abierto, con el fin avanzar hacia el objetivo general del proyecto: seleccionar un entorno virtual de enseñanza y aprendizaje (EVEA) a fin de que sea utilizado como entorno institucional de la UNTDF y contribuir al proceso de implementación.

Como resultado inicial se seleccionaron, en una primera etapa las Plataformas: Chamilo, Moodle, Ilias y Sakai.

Previo a la evaluación, se

establecieron y se aplicaron un conjunto de criterios basados en las funcionalidades provistas por cada EVEA. Se consideraron tanto aspectos pedagógicos como técnicos.

Con el fin de ir profundizando el análisis, se instalaron en el servidor de la Universidad, en un entorno de prueba, las cuatro plataformas mencionadas, asignando en cada una roles de Administrador, Docente y Alumno a los distintos integrantes del proyecto.

Finalmente, atendiendo a sus funcionalidades, y fundamentalmente a la simplicidad que presenta su uso para los docentes y los alumnos, hemos decidido seleccionar para las experiencias piloto previstas en el proyecto la plataforma Chamilo [11]. Durante este cuatrimestre se la utilizará en aquellas materias de 1º, 2º y 3º año de la Licenciatura en Sistemas que están a cargo de los docentes que participan del proyecto.

A partir de las experiencias obtenidas restará:

- Recomendar un entorno, acompañado de una fundamentación relacionada con el modelo de enseñanza y sus prestaciones como herramienta de mediación educativa.
- Indagar sobre las prácticas de enseñanza del cuerpo académico de la UNTDF, y en particular sobre sus conocimientos de TICs y el modo en que las aplican.
- Sugerir un plan de actualización de conocimientos para su uso.
- Implementar instancias de capacitación.
- Brindar soporte a un conjunto de experiencias de aplicación.

3. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo está formado por docentes - investigadores, de los

cuales uno (1) es Magister en Política y Gestión de la Ciencia y la Tecnología, dos (2) son Especialistas en Docencia Universitaria, tres (3) son Licenciados en Informática, uno (1) es Analista Programador Universitario y uno (1) es alumno avanzado de la carrera Licenciatura en Sistemas.

La Mg. Zangara ha entregado su tesis para acceder al Doctorado en Ciencias Informáticas en la Facultad de Informática de la UNLP. Se encuentra a la espera de la defensa.

El Licenciado Romano se encuentra cursando la Maestría en Tecnología Informática Aplicada en Educación en la Facultad de Informática de la UNLP.

El Licenciado Pendentí está desarrollando la tesis de la maestría en Ingeniería de Software en la Facultad de Informática de la UNLP.

El alumno Rojas está desarrollando su tesis de grado, bajo la dirección de las docentes Depetris y Zangara, sobre temas afines al proyecto.

Uno de los beneficios del proyecto es la consolidación de parte del grupo de I/D, que podrá asesorar y capacitar al personal docente en la implementación del EVEA seleccionado.

REFERENCIAS

[1] UNTDF. (2012, Accedido el 3 de marzo de 2017). Estatuto de la UNTDF [Online]. Disponible en www.untdf.edu.ar/reglamentacion.

[2] Patricia Avila M., Martha Diana Bosco H.. (2001, Accedido el 27 de febrero de 2017). Virtual Environment for Learning, A New Experience [Online]. Disponible en http://investigacion.ilce.edu.mx/panel_control/doc/c37ambientes.pdf.

[3] Ariel Ferreira Szpiniak, Cecilia V. Sanz. (2009, Marzo 2017) Hacia un modelo de evaluación de entornos virtuales de enseñanza y aprendizaje. La

importancia de la usabilidad. Revista TEyET [Online] N° 4 pp. 10 - 14. Disponible en: <http://teyet-revista.info.unlp.edu.ar/numero-4-octubre-de-2009/>

[4] M. A. López Carrasco, Aprendizaje, competencias y TIC (eBook). Pearson, 2012.

[6] Alvarez Diego. Plataformas de enseñanza virtual libres y sus características de extensión: Desarrollo de un bloque para la gestión de tutorías en Moodle. Tesis de Ingeniería. I.I. Universidad de Alcalá. Alcalá de Henares, España. Disponible en <http://www3.uah.es/libretics/files/Tutorias.pdf>. Accedido el 14 de octubre de 2016

[7] Baumgartner P, Häfele H., et al. E-Learning Praxishandbuch: Auswahl von ernplattformen. Marktübersicht - Funktionen Fachbegriffe. Innsbruck-Wien, StudienVerlag. 2002.

[8] Hernández Schäfer (2002, marzo 2017) L.E. Estado actual y futuro de las plataformas e-learning: estándares y especificaciones IMS. Universidad Católica del Norte, Chile. [Online] Disponible en <http://repositorial.cuaed.unam.mx:8080/jspui/bitstream/123456789/2974/1/1-cl-Luis%20Hern%C3%A1ndez%20Sc%C3%A4fer-Gim%20Ivy.pdf>

[9] Adell J. (2004, marzo 2017) Selección de un entorno virtual de enseñanza/aprendizaje de código fuente abierto para la Universitat Jaume I. Centre d'Educació i Noves Tecnologies de la UJI con la colaboración del Servei d'Informàtica y del Gabinet Tècnic del Rectorat. [Online]. Disponible en:

http://cent.uji.es/doc/eveauji_es.pdf

[10] Clarenc, C. A. Instrumento de evaluación y selección de sistemas de gestión de aprendizaje y otros materiales digitales: Medición y ponderación de LMS y CLMS, recursos

educativos digitales y herramientas o sitios de la WEB 3.0. Presentado en Congreso Virtual Mundial de e-Learning: Grupo GEIPITE. Octubre 2013

Disponible en scribd:

<http://es.scribd.com/doc/175057118/Instrumentoevaluacion->. Accedido el 17 de marzo de 2017.

[11] Plataforma Chamilo LMS: ¡Las novedades de 1.11! - Slideshare

Disponible en

<https://es.slideshare.net/chamiluda/plataforma-chamilo-lms-las-novedades-de-111>

Marzo 2018

Diseño y Aplicación de Estrategias para la Enseñanza Inicial de la Programación

Depetris Beatriz, Feierherd Guillermo, Pendenti Horacio, Aguil Mallea Daniel, Tejero Carlos Germán, Prisching Guillermo, Fierro Ariel, Aguilar Santiago, Domínguez Juan, Mamani Jonatan

Instituto de Desarrollo Económico e Innovación
Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur
Hipólito Irigoyen 880 - Ushuaia - Tierra del Fuego
{bdepetris, gfeierherd, hpendenti, daguilmallea, ctejero, gprisching}@untdf.edu.ar
arielalejandrofierro@gmail.com, santeex@gmail.com, juani.dz@hotmail.com,
jonush88@gmail.com

RESUMEN

La enseñanza y el aprendizaje inicial de la programación presentan importantes desafíos para los docentes y los alumnos de las carreras que requieren incorporarlos. Lejos de disminuir, estos han ido aumentando, como consecuencia de la necesidad de incorporar, a los conceptos tradicionales de programación, los que requiere la programación concurrente y paralela. [1]

La problemática no reside en la incapacidad para resolver problemas propiamente dicha sino en el escaso desarrollo del pensamiento computacional y de los procesos de abstracción que ellos requieren. [2].

El objetivo de esta investigación es mejorar algunas estrategias ya utilizadas y promover nuevas propuestas didácticas, que permitan afrontar dichos desafíos, buscando mejorar el desempeño académico de los alumnos ingresantes a las carreras de sistemas de la UNTDF.

Se prevé que la investigación propuesta habilite a encontrar nuevas dimensiones de análisis acerca de las prácticas docentes que aporten conocimiento significativo al campo de la didáctica de la programación. Además, podrían encontrarse propuestas adecuadas para la introducción de estos temas en los últimos años de la currícula en el nivel

secundario, no con el objetivo de que todos los alumnos se conviertan en futuros informáticos, sino que desarrollen su capacidad para resolver problemas.

Palabras clave:

Didáctica de la Programación, Robótica Educativa.

CONTEXTO

El presente proyecto forma parte del “Área de Investigación 8 - Desarrollo Informático”, del Instituto de Desarrollo Económico e Innovación de la Universidad Nacional de Tierra del Fuego (UNTDF).

El proyecto se presentó a la convocatoria realizada por la UNTDF en septiembre de 2016 y resultó aprobado luego de haber sido sometido a evaluación externa. (Resolución Rectoral 060/2017 del 10/04/2017).

El financiamiento requerido es suministrado por la UNTDF.

El proyecto se desarrolla formalmente desde el 01/03/2017 hasta el 28/02/2019.

INTRODUCCIÓN

La enseñanza y el aprendizaje de la programación de computadoras representa una de las principales dificultades que enfrentan los alumnos

que deciden estudiar carreras de la disciplina informática [3]. Esto se evidencia, al menos en las universidades argentinas, en elevados índices de fracaso en las materias iniciales que tratan estos temas.

Ahondando en las causas del fracaso, se ha concluido que las mismas no residen en la dificultad de los alumnos para traducir la solución de un problema a las sentencias propias de un lenguaje de programación, sino que, por el contrario, las mismas son mucho más profundas y tienen que ver con la falta de metodología, hábito y capacidad para resolver problemas. A su vez, se nota un marcado déficit de la capacidad de abstracción, la que, según Piaget, debería desarrollarse durante la adolescencia [4]. Por otra parte, es probable que, anticipándose a estas dificultades, muchos alumnos desistan de ingresar a este tipo de carreras luego de haber finalizado sus estudios secundarios.

No obstante, en la sociedad del conocimiento en la que estamos inmersos, las habilidades para resolver problemas son imprescindibles para desempeñarse en forma adecuada en cualquier ámbito. Las motivaciones del proyecto son, entonces, reducir los índices de deserción de los primeros años de las carreras universitarias de sistemas y, en la medida que el proyecto pueda alcanzar alumnos del nivel medio, contribuir a aumentar la matrícula en las mismas e incrementar las capacidades de estos para resolver problemas.

En el convencimiento de que las dificultades que se enfrentan tienen como causa el déficit en la capacidad de abstracción, se busca subsanarlo recurriendo a estrategias que permiten que los alumnos evolucionen hacia el pensamiento abstracto partiendo de experiencias concretas. [5]

En ese camino, el equipo de trabajo

ha desarrollado y viene utilizando con cierto éxito herramientas que permiten visualizar la ejecución de los algoritmos. A partir de desarrollos realizados en el Instituto de Investigación en Informática LIDI, de la Facultad de Informática de la UNLP, se evolucionaron los mismos, incorporando funcionalidades que permiten utilizarlos para la enseñanza inicial de la programación concurrente y paralela [4].

Por otra parte, en los últimos años muchas instituciones han recurrido al empleo de robots con fines educativos. El uso de robots como medio didáctico es una alternativa significativa como motivación para materias introductorias y avanzadas en carreras de Informática e Ingeniería Electrónica, y sirve como eje de integración horizontal y vertical de contenidos curriculares. [6] [7]

La robótica educativa es uno de los campos de investigación que mayor auge está teniendo como investigación aplicada.

En este proyecto se recurrirá al uso de robots didácticos de tecnología sencilla y al alcance de los presupuestos de una escuela, centro educativo o universidad de recursos limitados.

El proyecto incluye la adaptación de un robot (hardware) y del soporte informático (software) que permita su manejo por usuarios sin experiencia o formación específica. Mediante técnicas de enseñanza experimental el alumno inicial aprende nociones formales de algorítmica. Para el alumno avanzado de ingeniería y de las carreras de informática el robot proporciona la posibilidad de experimentar en forma práctica con conceptos avanzados. El robot se telecomanda desde una computadora personal y puede incorporar un sistema de sensado de posición y de obstáculos basado en el uso de sensores de colores y de distancia. Adicionalmente se podría

añadir la captura de imágenes mediante una cámara digital de bajo costo (tipo webcam). El alumno proporciona instrucciones en un lenguaje de alto nivel específicamente creado para introducirlo en conceptos relacionado con la disciplina de programación.

Por otra parte, un proyecto de estas características, dado el atractivo social que genera el uso de la tecnología y el muy bajo costo de la misma, es ideal para ser promocionado en colegios y otras instituciones educativas, buscando fomentar el interés en los estudios universitarios. Ello conseguirá atraer más y mejores alumnos a las carreras técnicas de la universidad.

A su vez, el uso de esta tecnología como herramienta de estudio en las materias de las carreras de informática, motivará a los alumnos de las mismas para participar en las asignaturas y así disminuir el índice de deserción, minimizar el tiempo de finalizado de la carrera y extender las experiencias más allá de lo meramente formal. El desarrollo es lo suficientemente genérico como para que puedan participar alumnos de diferentes carreras y diferentes niveles dentro de las mismas, favoreciendo la aplicación de los contenidos de cada materia en la solución de problemas concretos. Esto permite abordar un tema de gran importancia para la formación profesional, como es la integración curricular horizontal y vertical de los contenidos de varias asignaturas, lo cual hace posible que los alumnos comprendan el alcance y utilidad de los diferentes bloques curriculares y mejoren conceptualmente su articulación.

Varios han sido los grupos de investigación que han estudiado la utilización de la robótica desde la perspectiva docente, destacando los beneficios en los estudiantes en cuanto al desarrollo de habilidades tales como

creatividad, trabajo en equipo, autoaprendizaje e investigación y como facilitador del aprendizaje de contenidos teóricos en la disciplina de programación.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- a) La evolución de productos educativos ya desarrollados en proyectos anteriores, que integren la concurrencia y el paralelismo, poniendo énfasis en la visualización de la ejecución de los algoritmos.
- b) La inclusión de robótica educativa y su integración a los productos mencionados en a), tanto en materias iniciales de programación como en talleres de iniciación a la programación concurrente.

RESULTADOS OBTENIDOS / ESPERADOS

El proyecto se inició el 1° de marzo de 2017.

Hasta la fecha de esta presentación, se han realizado las siguientes tareas:

- a) Selección de robots: En función de sus características de simplicidad y flexibilidad se seleccionaron los siguientes:
 - Basados en placas Arduino
 - Lego MindStorms
- b) Integración con la aplicación ya existente “Davinci Concurrente”. Para ello se desarrolló, para el Robot Lego MindStorms, un intérprete que reconoce un conjunto limitado de primitivas del lenguaje de programación permitiendo que el mismo se

- desplace sobre una ciudad física.
- a) Desarrollar, para el robot Arduino, un intérprete que reconozca un conjunto limitado de primitivas del lenguaje de programación “Davinci Concurrente” que permitan que el robot se desplace sobre una ciudad física.
 - b) Desarrollar, al menos para alguno de los robots que se están utilizando, un intérprete que permita que varios robots coordinen sus desplazamientos por la ciudad física.
 - c) Organizar talleres para alumnos del nivel medio, buscando despertar en los jóvenes la motivación necesaria para introducirse tempranamente en el mundo de la programación.
 - d) Organizar talleres con alumnos de 1er. y 2do. año de las carreras de sistemas, a fin de despertar en ellos el interés para trabajar e investigar en el tema, como una forma más de ayudar a disminuir la deserción en los primeros años. Se parte de la idea de que la motivación es un elemento indiscutible para mejorar los rendimientos académicos.

FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo está formado por docentes - investigadores, de los cuales dos (2) son Especialistas en Docencia Universitaria, cuatro (4) son Licenciados en Informática, uno (1) es Ingeniero en Sistemas de Computación y tres (3) son alumnos avanzados de la carrera Licenciatura en Sistemas.

Los Licenciados Pendentí y Aguil Mallea están desarrollando su tesis de Maestría en Ingeniería de Software en la Facultad de Informática de la UNLP.

El alumno Mamaní ha presentado para su defensa su tesis de grado, bajo la dirección de los docentes Depetris y Feierherd sobre temas afines al proyecto.

REFERENCIAS

- [1] Depetris B., Feierherd G. y otros. “Experiencias con Da Vinci Concurrente en la enseñanza inicial de la programación y la programación concurrente.” En Proceedings del VIII Congreso de Tecnología en Educación y Educación en Tecnología. Disponible en <http://sedici.unlp.edu.ar/handle/10915/27581> Accedido el 3 de marzo de 2018.
- [2] Berrocoso J., Fernández Sánchez M. R., Garrido Arroyo M.; “El pensamiento computacional y las nuevas ecologías del aprendizaje”; RED - Revista de Educación a Distancia, Septiembre de 2015. Disponible en http://www.um.es/ead/red/46/valverde_et_al.pdf. Accedido el 3 de marzo de 2017.
- [3] Blake, J.D.; “Language considerations in the first year CS curriculum.” J. Comput. Sci. Coll. 26, 6 (2011), 124–129.
- [4] Depetris B., Feierherd G., Aguil Mallea D., Tejero G.; “A multiplatform interpreter to introduce structured and concurrent programming”; Computer Science & Technology Series. RedUNCI; ISBN 978-987- 1985- 20-3; Año 2013
- [5] Cátedra Conceptos de Algoritmos, Datos y Programas, UNLP; “Por qué ‘pensar algoritmos’ es tan importante en Informática?” Revista Bit & Byte, Año 02, Número 04, diciembre 2016; Disponible en: <http://sedici.unlp.edu.ar/handle/10915/57362>. Accedido el 2 de marzo de 2018.
- [6] Pásztor A., Pap-Szigeti R., Török E.; “Mobile Robots in Teaching Programming for IT Engineers and its Effects”. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 11, 2013.

Disponible en <https://pdfs.semanticscholar.org/15af/a68042124356ed959587bf657783c1e70e8f.pdf>. Accedido el 10 de marzo de 2018.

[7] Correa C., Ferreira Szpiniak A.; “Laboratorios Teórico-Prácticos en Robótica Educativa”; XI Congreso de Educación en Tecnología y Tecnología en Educación. 2016. Disponible en http://sedici.unlp.edu.ar/bitstream/handle/10915/53812/Documento_completo.pdf-PDFA.pdf?sequence=1. Accedido el 2 de marzo de 2018

REFORTICCA: Recursos para el Empoderamiento de FORMadores en TIC, Ciencias y Ambiente

María José Abásolo^{1,2}, Cecilia Sanz¹, Graciela Santos³, María L. Castro³, Andrea Miranda³, Gabriela Cenich³, María José Bouciguez³, Cecilia Papini³,

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata (UNLP)
{mjabasolo, csanz}@lidi.info.unlp.edu.ar

²Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA)

³Centro de Educación en Ciencias con Tecnología (ECienTec)
Facultad de Ciencias Exactas - Universidad Nacional del Centro Pcia. Bs. As.
{nsantos, mlcastro, amiranda, gabcen, mjbouci}@exa.unicen.edu.ar

Resumen

La línea de investigación y desarrollo que aquí se presenta tiene por objetivo brindar a docentes de educación primaria y secundaria, y a la comunidad educativa en general, recursos educativos y herramientas digitales para ser usadas como soporte del proceso de enseñanza y aprendizaje de ciencias físicas, matemáticas, medio ambiente, y aportar al desarrollo del pensamiento computacional. En particular, se estudian, desarrollan y evalúan aplicaciones de realidad virtual, realidad aumentada, simulaciones, juegos e interacción tangible, con el objeto de resolver procesos de enseñanza y aprendizaje. Además, se propone crear una comunidad virtual de docentes para la construcción colaborativa de propuestas de enseñanza.

Palabras Clave: Realidad Virtual, Realidad Aumentada, Interacción Tangible, Enseñanza, Aprendizaje, Formación Docente, TIC

Contexto

En esta línea se lleva a cabo el proyecto “REFORTICCA Recursos para el Empoderamiento de FORMadores en TIC, Ciencias y Ambiente” aprobado como

Proyecto de Innovación y Transferencia en Áreas Prioritarias de la Provincia de

Buenos Aires de la CICPBA, el cual se lleva a cabo entre los grupos de investigación III-LIDI (UNLP), ECienTec y CIFICEN (UNICEN) de las diferentes áreas involucradas: TICs aplicadas a educación en ciencias físicas, matemática y medio ambiente.

Introducción

Los jóvenes y niños, nativos en la era digital, tienen un acercamiento natural a las Tecnologías de Información y Comunicación (TIC). Por otra parte es conocido que los docentes aún hacen uso escaso o insuficiente de la tecnología en la enseñanza, y en particular para la enseñanza de Matemática, Física o Ambiente (Cenich et al, 2017) (Arriasecq y Santos, 2017). Se asume que esto se debe a dos razones de peso: por un lado, desconocimiento de las nuevas tecnologías y, por otro, la falta de formación para el diseño de propuestas de enseñanza que integren las TIC. Se parte del supuesto que son los docentes quienes deben apropiarse de la tecnología y adquirir las competencias necesarias para enriquecer sus prácticas docentes con la tecnología para incidir en una mejora de la educación en ciencias en el ámbito de la Provincia de Buenos Aires. Es de central importancia el papel del docente responsable del diseño de las propuestas pedagógicas, para guiar

el proceso de enseñanza y aprendizaje, decidir qué herramienta será más apropiada, así como anticipar dificultades y obstáculos de aprendizaje.

En los últimos años han surgido nuevas tecnologías y paradigmas como la Realidad Aumentada (RA), la Realidad Virtual (RV), la Interacción Tangible (IT), que se presentan como herramientas que pueden apoyar los procesos de enseñanza y aprendizaje tanto en los diferentes niveles educativos de la educación formal como en educación no formal, y educación especial (Hernández Ortega et al., 2012).

El objetivo general de investigación es brindar recursos educativos y herramientas TIC -basadas en RA, RV, simulaciones, juegos e IT- para ser usadas como soporte del proceso de enseñanza y aprendizaje de ciencias físicas, matemáticas y medio ambiente. En particular se orienta a empoderar a los docentes mediante el desarrollo de competencias TIC, ofreciendo cursos de capacitación sobre el uso de las TIC en la enseñanza, poniendo a disposición recursos que integren los conocimientos disciplinares, didácticos y tecnológicos. Se propone crear una comunidad de docentes para la construcción colaborativa de Objetos de Enseñanza (OE) posibilitando el intercambio de experiencias de aula donde se realice integración curricular de TIC. Los OE son propuestas didácticas creadas de manera conjunta entre docentes que pueden ser adaptadas e implementadas en el aula. Los docentes pueden sugerir modificaciones que serán consideradas en un proceso de revisión de tales objetos a cargo de la comunidad virtual. Esta posibilidad permitirá incrementar la cantidad de recursos para el aula que contemplan diferentes puntos de vista sobre un mismo objeto de conocimiento.

Realidad Aumentada (RA)

La RA propone aumentar las capacidades de percepción humana, para visualizar la información digital - como texto, imágenes, vídeos o animaciones en 3D-

directamente embebida en el mundo físico real (Manresa-Yee et al, 2011).

Las primeras aplicaciones de RA exigían la utilización de dispositivos específicos como gafas de realidad aumentada con seguimiento (tracking) del usuario. El uso de RA comenzó a conocerse a partir de aplicaciones de escritorio y aplicaciones web que solo requerían una cámara web y monitor. Recientemente la RA se popularizó con el uso de aplicaciones móviles que se ejecutan en tablets o celulares, las cuales utilizan diferentes formas aumentar la realidad:

- Captación de códigos QRCode y códigos de barra, para mostrar información relacionada o visitar sitios web
- Captación de imágenes conocidas, como los marcadores de RA o imágenes en general, para mostrar información y modelos 3D registrados espacialmente con los marcadores
- Geolocalización del participante, es decir, la localización de su ubicación en el globo terráqueo, en base a dispositivos físicos como GPS, brújulas, acelerómetros y giroscopios incorporados a celulares de última generación. Se provee contenidos relevantes del entorno - indicadores de puntos de interés (POIs)- dependiendo de la ubicación del usuario en un cierto lugar en el mapa terrestre.

Existe una variedad de aplicaciones de RA que pueden encontrarse en la web para apoyar el proceso de enseñanza y aprendizaje en áreas como biología, matemática, astronomía, anatomía, idiomas, ecología, etc. Existe una creciente cantidad de estudios publicados que reportan ventajas, limitaciones, desafíos, etc. de la RA en educación (Gavilanes et al, 2018). Sin embargo, dado que la RA es una tecnología emergente, es importante revisar los avances y el impacto real en escenarios educativos.

Realidad Virtual (RV), Simulaciones y Juegos

RV es un término que se aplica a experiencias visuales donde el participante se ve inmerso e interactúa en un ambiente o escena virtual 3D con diferentes grados de inmersión (Manresa-Yee et al, 2011). Si bien existen diferentes dispositivos de visualización utilizados en RV, la tecnología más popularizadas son las gafas de vídeo. El precio de estos dispositivos era una de las limitaciones para su uso en educación, hasta que recientemente comenzó a popularizarse el uso de celulares convertidos a gafas mediante el uso de soportes de cartón (Google Cardboard¹). Entre las aplicaciones educativas se resalta la realización de visitas virtuales a lugares que, por su situación geográfica, no se pueden visitar de forma habitual. Como ejemplo puede citarse Google Expeditions² es una aplicación para Android, que permite explorar sitios de interés desde el salón de clase utilizando las gafas Google Cardboard.

Por su parte, los videojuegos y simulaciones de procesos y fenómenos se presentan como entornos de aprendizaje interactivo participativo que cautiva a un jugador ofreciendo desafíos que requieren mayores niveles de dominio, que pueden conjugar la participación y diversión con el rigor y la resolución de nuevas situaciones (Aldrich, 2009) (Squire, 2008). Los videojuegos y simulaciones se emplean en la enseñanza de las ciencias. (Boucíguez et al, 2013) (Santos, 2016) (Paoletti et al, 2017) describen algunos videojuegos educativos para enseñar Física. Puede citarse GeoGebra como uno de los software para el modelado y simulación en Matemática y en Física seleccionado para trabajar, dada su versatilidad y potencialidad (Martinovic, 2014) (Sezen Yüksel y Çıldır, 2015). Cabe destacar que es un software de uso y distribución gratuita y que su potencialidad educativa yace en que permite relacionar distintas

representaciones de un problema (Papini y Miranda, 2016).

Interacción Tangible (IT)

La IT permite, a través de la manipulación física de objetos superficies y espacios del entorno real, denominados representaciones tangibles, controlar sus contrapartes digitales. La IT ofrece oportunidades para el desarrollo de actividades educativas colaborativas, dado que permite que el grupo se concentre en el objetivo a cumplir y la tecnología no resulte un distractor (Artola, 2013).

Líneas de investigación y desarrollo

- RA aplicada en el salón de clases y en exposiciones de ciencia
- RV aplicada en el salón de clases y en exposiciones de ciencia
- Simulaciones y Videojuegos educativos para la enseñanza de Matemática y Física
- IT aplicada a la educación
- Acercamiento de los docentes a las tecnologías emergentes

Resultados y Objetivos

Durante el año 2017 se ha avanzado en los siguientes resultados y objetivos

- Se realizó el taller “La integración de TICs en el aula”, por parte de docentes de ECienTec UNICEN, Tandil, nov 2017. Asistieron 15 docentes provenientes de diferentes localidades de la Provincia de Bs.As: Olavarría, Hinojo, Laprida, Ayacucho, Bolívar, Azul y Tandil. Se abordaron aspectos metodológicos y prácticos sobre la integración de TIC en escenarios educativos. Este año se continuará con esta dinámica de encuentros y se extenderá a otros docentes y escuelas de la provincia.
- Se formó una comunidad virtual de 27 docentes para intercambiar experiencias y abordar en colaboración la producción de OE. Como parte de una tesis de

¹ <https://vr.google.com/cardboard/>

² <https://edu.google.com/expeditions/>

grado se está desarrollando un espacio virtual que permite a los docentes integrantes de la comunidad virtual de aprendizaje la construcción colaborativa de propuestas de OE.

- Se elaboró una encuesta para indagar sobre conocimiento disciplinar, pedagógico y tecnológico de docentes de ciencia (matemática, física, tecnología, medio ambiente) del Centro de la provincia Buenos Aires. Se recolectaron datos recolectados de docentes en diferentes encuentros mediante el llenado del formulario diseñado en Google. Se planea continuar con la toma de datos y se espera recolectar información sobre aproximadamente 200 docentes.
- Se crearon diferentes materiales educativos digitales e impresos. En particular, se difundió un folleto sobre el concepto de autorregulación orientado a fomentar estrategias de estudio en estudiantes cercanos a ingresar a la universidad (Sanz & Artola, 2017).
- Se desarrolló un juego de IT para la enseñanza de fracciones que fue presentando en la Exposición de la UNLP para presentar las ofertas de carreras de esta institución. También se presentó en las jornadas de ciencia y tecnología de la Facultad de Informática (Nordio, Artola & Sanz, 2017).
- Se diseñaron ejemplos de prácticas con RA y se presentaron en el marco de un curso de postgrado de la Facultad de Informática de la UNLP. En particular se diseñó una actividad para la relación entre pintores de diferentes corrientes y sus obras.
- Se trabajó en la difusión de estrategias para el seguimiento de actividades colaborativas mediadas por TIC y se publicó un trabajo (Zangara & Sanz, 2017).
- Se incorporaron actividades utilizando RV con celulares y Google Cardboard en las exposiciones interactivas de ciencias realizadas durante la Semana

Nacional de la Ciencia y la Tecnología (Campus Universitario UNICEN, septiembre 2017) y en el Parador Cultural en la Playa, actividad de verano organizada por la Secretaría de Cultura de UNICEN (Quequén, enero 2018).

- Se dictó el taller de ciencias “Animaciones en Geogebra: una herramienta para estudiar Matemática” 17.º Foro de Enseñanza de Ciencias y Tecnologías, mayo de 2017, CABA.
- Se está desarrollando el Objeto de Aprendizaje, siguiendo metodología CROA³, “Carros de Laboratorio del Plan Nacional Ciencias Naturales” del Ministerio de Educación, que tiene por objeto dar a conocer a los docentes los elementos que componen un kit de química distribuido en las escuelas primarias. Mediante la captación de un código QR pegado a cada elemento del kit se accederá a la información asociada.
- Se está desarrollando como tesis de grado una aplicación móvil para el consumo responsable de productos alimenticios y de limpieza.

Formación de recursos humanos

La formación de recursos humanos es prioritaria en esta línea, y por esto se implementó la carrera de postgrado mencionada en la sección anterior. En el marco de los proyectos conjuntos mencionados se ha podido contar con la visita de profesores de otras universidades nacionales y extranjeras (UNLP, UNS, UNICEN y UIB), con experiencia y formación en los temas propuestos, los cuales impartieron cursos de posgrado.

En la actualidad hay en curso diferentes tesis de postgrado, grado y becas de entrenamiento en el marco de esta línea de investigación:

- Mario Vincenzi. “La Realidad Aumentada en la educación. Vigencia,

³ <http://croa.info.unlp.edu.ar>

proyecciones y límites” Director: Abásolo, M.J. (tesis de especialización en curso).

- Natalia Encina. “Evaluación de browsers de realidad aumentada para apoyar procesos de enseñanza - aprendizaje” Director: Abásolo, M.J. (tesis de especialización en curso).
- Wilma Gavilanes “Metodología para la evaluación del impacto de experiencias con Realidad aumentada en educación superior” Director: Abásolo, M.J. (tesis de doctorado en curso).
- María José Bouciguez “Ambientes virtuales altamente interactivos basados en videojuegos y simulaciones para la educación en ciencias” Directores: Santos, G. y Abásolo, M.J. (tesis de doctorado en curso)
- Milagros Paoletti. “Estrategias y conocimientos durante un juego educativo”. Becaria EVC-CIN 2016 Directores: Santos, G. y Miranda, A.
- Marisa Salerno “Herramienta colaborativa para la creación de Objetos de Enseñanza” Directores: Miranda, A. y Jonás, I. (tesis de grado en curso).

Referencias

Aldrich, C. (2009). *Virtual worlds, simulations, and games for education: A unifying view*. Innovate 5 (5), http://www.innovateonline.info/pdf/vol5_issue5/Virtual_Worlds,_Simulations,_and_Games_for_Education- A_Unifying_View.pdf

Irene Arriaseq; Graciela Santos (2017) “Nuevas tecnologías de la información como facilitadoras de Aprendizaje significativo”. Revista Archivos de Ciencias de la Educación. La Plata: Departamento de Ciencias de la Educación, Facultad de Humanidades y Ciencias de la Educación, Universidad Nacional de La Plata. 2017 vol.12 n°12. p - .ISSN 2346-8866

Artola, Verónica; Sanz, Cecilia; Moralejo, Lucrecia; Pesado, Patricia; Baldassarri, Sandra (2015) *Herramienta de autor para la creación de actividades educativas basadas en Interacción Tangible*. XIII Workshop de Tecnología Informática Aplicada en Educación – CACIC 2015. Junín, Proceeding del Congreso. ISBN: 978-987-3724-37-4 Buenos Aires, Argentina. Octubre 2015.

Artola, Verónica (2013) *Diseño e implementación de un prototipo basado en este paradigma de interacción orientado al aprendizaje colaborativo*. Tesis de grado Facultad de Informática UNLP <http://sedici.unlp.edu.ar/handle/10915/46826>

Adams Becker, S., Cummins, M., Davis, A., and Yuhnke, B. (2016). *2016 NMC Technology Outlook for Australian Tertiary Education: A Horizon Project* Regional Report. Austin, Texas: The New Media Consortium

María José Bouciguez, Graciela Santos, María José Abásolo (2013) *Potencialidad de los videojuegos en el aprendizaje de Física*. Actas de WEFA 2013 I Workshop de Enseñanza de Física en Argentina, Universidad Nacional del Centro de la Pcia. de Bs. As., ISBN 978-950-658-342-2.

Gabriela Cenich; Sonia Araujo; Graciela Santos (2017) “TIC y culturas de enseñanza. Elaboración de una encuesta para indagar los usos educativos de las TIC por docentes de Matemática” Revista Iberoamericana de Educación. Madrid: Centro de Altos Estudios Universitarios (CAEU-OEI). vol.73 n°1. p9 - 28. issn 1022-6508. E-ISSN 1681-5653

Coll, C. Muri, M. T. y Onrubia, J. (2008). *Análisis de los usos reales de las TIC en contextos educativos formales: una aproximación socio-cultural*. En Revista Electrónica de Investigación Educativa, 10(1). Consultado el 28 de agosto de 2012, en: <http://redie.uabc.mx/vol10no1/contenido-coll2.html>

Gavilanes, W., Abásolo, M.J., Cuji, B. (2018) “Realidad Aumentada en la Educación: una Revisión desde la Perspectiva Pedagógica”, Revista ESPACIOS. ISSN 0798 1015, en prensa

Hernández Ortega, José; Pennesi Fruscio, Massimo; Sobrino López, Diego y Vázquez Gutiérrez, Azucena (2012) *Tendencias Emergentes en Educación con TIC*. Asociación Espiral, Educación y Tecnología, ISBN: 978-84-616-0448-7

Manresa-Yee, C.; Abásolo, M.J.; Mas Sansó, R.; Vénere, M.. (2011) *Realidad Virtual, Realidad Aumentada e Interfaces Basadas en Vision*. XV Escuela Internacional de Informática, XVII Congreso Argentino de Ciencia de la Computación CACIC 2011. Editorial EDULP, ISBN 978-950-34-0765-3

Martinovic, D., Karadag, Z., & McDougall, D. (Eds.) (2014). *Proceedings of the Fifth North American GeoGebra Conference*, GeoGebra-NA 2014, November 21-22, 2014, Toronto, ON: University of Toronto.

Moralejo L., Sanz C., Pesado P., Baldassarri S. (2013) *AuthorAR: Authoring Tool For Building*

Educational Activities Based On Augmented Reality
International Conference on Collaboration
Technologies and Systems (CTS 2013) San Diego,
Estados Unidos. Mayo de 2013. Proceeding del
Congreso. Páginas 377-381. ISBN: 978-1-4673-
6404-1

Nordio, M.; Artola, V.; Sanz, C. (2017).
*FraccionAR: juego sobre fracciones basado en
Interacción Tangible*. Disponible en:
[http://weblidi.info.unlp.edu.ar/catedras/proyectoAlu/
video.mp4](http://weblidi.info.unlp.edu.ar/catedras/proyectoAlu/video.mp4)

M. Paoletti; D, García; A, Miranda; G, Santos
(2017) *Conocimiento en acción: una propuesta
para aprender choque a partir de un videojuego*.
Argentina. Concordia. Congreso. REF XX. APFA,
Facultad de Ciencias de la Alimentación de la
Universidad Nacional de Entre Ríos, la Facultad
Regional Concordia de la Universidad Tecnológica
Nacional y el Instituto Superior de Disciplinas
Industriales y Ciencias Agropecuarias

Papini, C. y Miranda, A. (2016) *Análisis didáctico
de un problema matemático para una clase de
secundaria en la que se utiliza el programa
Geogebra*. Cap. 1 del libro: “Pasaporte a la
enseñanza de las ciencias. La modelización como
eje organizador para la construcción de
significados”, Consuelo Escudero y Silvia Stipcich
(comp.). Noveduc libros del Centro de
Publicaciones Educativas y Material Didáctico
S.R.L. En proceso de edición ISBN 978-987-538-
482-8.

Santos G. (2016) *Videojuegos y estrategias para
enseñar física*. Cap. 3 del libro: “Pasaporte a la
enseñanza de las ciencias. La modelización como
eje organizador para la construcción de
significados”, Consuelo Escudero y Silvia Stipcich
(comp.). Noveduc libros del Centro de
Publicaciones Educativas y Material Didáctico
S.R.L. En proceso de edición. 2016. ISBN 978-
987-538-482-8

Sezen Yüksel & Çıldır. (2015). *The Impacts of
Dynamic Geometry Software on Graphing Abilities
of Prospective Physics Teachers: GeoGebra
Sample*. Eurasian J. Phys. & Chem. Educ. 7(1): 46-
61.

Squire, K.D. (2008) *Game-based learning: An
emerging paradigm for learning*. Performance
Improvement Quarterly, 21 (2), 7-36.
[http://www3.interscience.wiley.com/journal/12083/
5177/issue](http://www3.interscience.wiley.com/journal/12083/5177/issue)

Sanz, C. & Artola, V. (2017) *Folleto de
Autorregulación*. Vídeo disponible en:
[http://weblidi.info.unlp.edu.ar/autorregulacion/vide
o1.mp4](http://weblidi.info.unlp.edu.ar/autorregulacion/video1.mp4)

Zangara, A. & Sanz, C. (2017) “*Displaying the
collaborative process as meta-knowledge.
Description of a mirroring strategy and its results*”.
En: Communications in Computer and Information
Science (CCIS) vol. 790. P. 79-89. Springer

M-learning con Realidad Aumentada basada en Objetos 3D

Susana I. Herrera¹, Cecilia V. Sanz², María I. Morales¹, Rosa A. Palavecino¹, Marilena Maldonado¹, Ivana Irurzun¹, Alvaro J. Carranza¹, Gabriela I. Suárez¹

¹*Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero*

²*Instituto de Investigación en Informática LIDI, Facultad de Informática, Universidad Nacional de La Plata*

{sherrera, rosypgg, marilena}@unse.edu.ar, csanz@lidi.info.unlp.edu.ar,
mines_morales@yahoo.com.ar, ivanairurzun@gmail.com, carranza1903@hotmail.com

Resumen

Esta investigación estudia alternativas para diseñar prácticas de “Aprendizaje basado en dispositivos móviles” (m-learning) enriquecidas con el uso de Realidad Aumentada (RA) con objetos 3D.

Las principales temáticas relacionadas con la investigación son: desarrollo de aplicaciones móviles multiplataforma, bibliotecas de funciones para la manipulación de objetos 3D, herramientas para la creación de objetos 3D, generación de repositorios de objetos 3D y bibliotecas de funciones para manipular repositorios en aplicaciones móviles.

A partir de necesidades pedagógicas observadas en la cátedra de Álgebra Lineal de la Facultad de Ciencias Exactas y Tecnologías (FCEyT) de la Universidad Nacional de Santiago del Estero (UNSE), fue diseñada la aplicación AlgeRA, que constituye el principal resultado de esta investigación. Actualmente, se encuentra en desarrollo.

Palabras clave: m-learning, realidad aumentada, aplicaciones móviles, objetos 3D, repositorio de objetos 3D.

1 Contexto

Esta investigación se lleva a cabo en el marco del proyecto denominado “Computación Móvil: desarrollo de aplicaciones y análisis forense”,

financiado por el Consejo de Ciencia y Técnica de la UNSE, durante el período enero 2017-diciembre 2018.

La investigación cuenta con el asesoramiento del Laboratorio de Investigaciones en Informática LIDI, de la Facultad de Informática de la Universidad Nacional de La Plata; y de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta. Asimismo, se trabaja en forma coordinada con el proyecto de investigación en Computación Móvil de la Universidad Católica de Santiago del Estero. Próximamente se firmará un convenio específico de colaboración en investigación entre UNSE y UCSE.

2 Introducción

El auge de los dispositivos móviles y la evolución de las redes móviles han impulsado el desarrollo del m-learning. Con el acelerado crecimiento del porcentaje de población mundial que lo utiliza ha surgido una nueva preocupación acerca de su uso: la calidad de los recursos que se elaboran y el impacto en la mejora del aprendizaje [1].

Entre las ventajas que se pueden encontrar al utilizar m-learning, se destacan: uso eficiente del tiempo, expansión de la alfabetización digital, accesibilidad, contacto social, mejoramiento de la productividad, aprendizaje colaborativo, incremento en

el estudio individual, información eficaz y los profesores pueden diseñar y poner a disposición de los alumnos materiales que contribuyan al aprendizaje de sus alumnos, pudiendo estos, disponer de dichos materiales de manera asíncrona [9].

En el marco de las investigaciones sobre computación móvil de UNSE, desde el año 2013, se desarrollaron dos aplicaciones móviles para aprendizaje: Educ-mobile [4] e Ima-Colab [8]. Asimismo, se diseñó y validó un marco para el análisis, diseño y evaluación de experiencias de m-learning, denominado MADE-mlearn [5]. Usando dicho marco, se diseñaron e implementaron diversas prácticas de m-learning en escuelas primarias rurales, en escuelas secundarias, en carreras de grado y posgrado, abarcando el aprendizaje en Tecnologías, Programación y Matemática [4, 8].

Con el propósito de optimizar los resultados de aprendizaje y la motivación en m-learning, esta investigación estudia una de las tecnologías emergentes con posibilidad de inclusión en educación: la RA. Esta se caracteriza por: (a) la combinación de objetos virtuales y reales en un escenario real, (b) usuarios interactuando en tiempo real y (c) una alineación entre los objetos reales y virtuales. Respecto a su potencialidad en el plano educativo, la RA permite incorporar multimedia a los procesos de enseñar y aprender, innovar en la práctica docente y promover el diseño de materiales educativos atendiendo a los requerimientos didácticos [3].

Integrar m-learning con RA para apoyar los procesos de aprendizaje y enseñanza, abre posibilidades para explorar nuevos enfoques y metodologías para la optimización y el fortalecimiento del aprendizaje.

Como lo señala [2], son numerosos los trabajos realizados usando las tecnologías mencionadas, aplicados a la enseñanza en distintas disciplinas desde ingeniería,

arquitectura, urbanismo, medicina, arte e historia, aprendizaje de idiomas, ciencias naturales, química, física y geografía.

La RA puede adoptar diferentes formas, permitiendo diversas posibilidades en el aula. Sus aplicaciones van desde la visualización de modelos 3D, a la incorporación de información adicional en recursos y materiales didácticos impresos o la creación de rutas geolocalizadas que permiten asociar información a lugares del entorno [10].

Entonces, la RA es atractiva, dado que se alinea con el aprendizaje activo [7, 6] recalca la importancia de los objetos 3D en el aprendizaje, ya que proveen: sensación de tocar, sentido, orientación y posición en el espacio. Permiten ver y experimentar información que es dinámica e interactiva.

Teniendo en cuenta lo planteado, esta investigación consiste en diseñar y desarrollar una práctica de m-learning para la enseñanza de Álgebra Lineal con RA basada en el uso de objetos 3D.

Si bien es notoria la importancia del uso de objetos 3D y RA en el aprendizaje, no es fácil la gestión y reutilización de estos objetos, al menos en el ámbito educativo. Según [1], si bien existen algunos repositorios de objetos 3D, no son de calidad o no presentan los recursos que se necesitan. Al mismo tiempo, es deseable contar con repositorios de objetos 3D educativos, que estén descriptos a partir de metadatos estandarizados para facilitar su almacenamiento, búsqueda y recuperación.

Si bien, existen repositorios en línea de objetos 3D que son accesibles y abiertos, por ejemplo:

- **Github**: es principalmente conocido por servir como un repositorio de código en línea.
- **Youmagine**: tiene una gran cantidad de modelos 3D, también brinda soporte a usuarios de impresoras 3D.
- **Yeggi**: es un sitio web que tiene como objetivo ser un motor de búsqueda/índice de modelos 3D.

- **3dwarehouse:** contiene millones de modelos creados en SketchUp, la aplicación de diseño y modelado de 3D más popular actualmente en el mundo.

3 Líneas de investigación y desarrollo

Esta investigación se organiza en torno a la siguiente línea:

- M-learning: Diseño de experiencias con RA para diferentes niveles educativos, diseño de repositorio de experiencias de m-learning y de repositorio de objetos 3D para RA.

Los objetivos son:

- Diseñar e implementar actividades de m-learning que involucren el uso de aplicaciones móviles multiplataforma basadas en RA, usando MADE-mlearn [10], tanto en nivel secundario como universitario.
- Diseñar un repositorio para las experiencias de m-learning basadas en MADE-mlearn.
- Analizar y clasificar herramientas (bibliotecas de funciones, herramientas de autor) para implementar RA en móviles en diversos formatos (texto, imágenes, objetos 3D).
- Caracterizar los objetos 3D que pueden ser usados en actividades educativas con RA en móviles.
- Diseñar un repositorio para la gestión de objetos 3D educativos para RA.

4 Resultados

En el primer año de investigación se obtuvieron los siguientes resultados parciales:

- Desarrollo del módulo de Gestión de Experiencias de la aplicación móvil ImaColab. Esto permite que la aplicación pueda ser usada de manera sencilla en diversas prácticas para diferentes niveles educativos.
- Rediseño, usando MADE-mlearn, de práctica de m-learning con ImaColab,

en nivel secundario, asignatura Programación.

- Implementación de prácticas de m-learning en nivel secundario.
- Diseño, usando MADE-mlearn, de práctica de m-learning para la enseñanza de Sistemas de Ecuaciones Lineales, para la asignatura Algebra Lineal, de las carreras de Ingeniería de la UNSE. La práctica consta de dos ejercicios. Uno de ellos consiste en mostrar, usando RA, un ejemplo de Sistemas de Ecuaciones Lineales de la vida cotidiana: Sistema de Riego por aspersión. El otro ejercicio se trata de un trabajo de campo que deben realizar los alumnos en forma grupal, recabando información usando el móvil, en puntos estratégicos de una red de tránsito urbano. El trabajo de campo consiste en medir, de manera simultánea, el caudal de vehículos. Luego, con dicha información, deben construir el sistema de ecuaciones lineales que modeliza la red de tránsito y las variaciones indicadas por el profesor en la consigna.
- Diseño de la aplicación AlgeRA, a partir del ítem anterior. La aplicación consiste en una aplicación móvil que tiene diversas funcionalidades o ejercicios relacionados con Algebra Lineal. Sin embargo, el prototipo sólo contendrá los ejercicios mencionados en el punto anterior. El diseño se hizo definiendo historias de usuario, utilizando Mobile-D.
- Análisis comparativos de herramientas para la construcción y manipulación de objetos 3D para RA.

Se espera terminar la aplicación AlgeRA para fines del primer semestre 2018 e implementar la práctica de m-learning usando dicha aplicación en el segundo semestre de 2018.

En relación a la aplicación ImaColab, será utilizada en asignaturas de Inglés (de la Facultad de Humanidades de UNSE) y en asignaturas de Tecnología del nivel

secundario en Cali, Colombia (en función de un acuerdo de colaboración con Unicatólica).

Además, se pretende trabajar en métodos para la construcción de Objetos de Aprendizaje para contextos móviles. Así como también en prácticas de m-learning con RA para niños con discapacidad auditiva.

5 Formación de recursos humanos

Esta investigación contribuye al desarrollo de una tesis de Especialización en Enseñanza en Tecnologías (UNSE), una tesis de Maestría en Informática Educativa (UNSE), dos trabajos finales de grado de la Licenciatura en Sistemas de Información (UNSE), una tesis doctoral (UNLP) y un posdoc sobre “m-learning con RA para niños con discapacidad auditiva” (UNSE-UNLP-Universidad Islas Baleares).

6 Referencias

1. Adair, M., Carina, F., Claudia, A., Susana, P., & Re. (2016). *Computación Aplicada: Búsqueda y Desarrollo de Nuevas Estrategias de*. XVIII Workshop de Investigadores en Ciencias de la Computación. Concordia.
2. Cabero Almenara, J., García Jiménez, F., Barroso Osuna, J. (2016). La producción de objetos de aprendizaje en “Realidad Aumentada”: la experiencia del SAV de la Universidad de Sevilla, *Revista Internacional de Investigación e Innovación Educativa, España*.
3. Gibelli, T., Graziani A. & Sanz C., (2017) Revisión de herramientas para la creación de modelos 3D orientados a la enseñanza de la matemática con realidad aumentada, XXIII Congreso Argentino de Ciencias de la Computación.
4. Herrera, S., Sanz, C. (2014). *Collaborative m-learning practice using Educ-Mobile*. Estados Unidos de América. Minneapolis. Libro. Artículo Completo. Conferencia. International Conference on Collaboration Technologies and Systems (CTS) 2014. University of Minnesota.
5. Herrera, S., Sanz, C., Fennema, C. (2013). MADE-mlearn: un marco para el análisis, diseño y evaluación de experiencias de m-learning en el nivel de postgrado. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*. La Plata: Universidad Nacional de La Plata. vol. n°10. p7 - 15. issn 1850-9959.
6. Krange, I., Fjuk, A., Larsen, A., Ludvigsen, S. (2002). Describing construction of. In *Proceedings of the Conference on Computer* (págs. 82-91). CSCL Community.
7. Moralejo, L. (2014). Análisis comparativo de herramientas de autor para la creación de actividades de realidad aumentada. La Plata.
8. Palavecino, R., Herrera, S., Sanz, C., Irurzun, I., Carranza, A. (2016). M-learning: aprendizaje de estructuras de datos con Ima-Colab. XI Congreso de Tecnología en Educación y Educación en Tecnología. RED UNCI Argentina. Morón. 2016.
9. Pedraza, L. E. & Valbuena S., D. (2014) M-learning y realidad aumentada, tecnologías integradas para apoyar la enseñanza del cálculo, *Revista de investigaciones UNAD*, Volumen 13. Número 2.
10. Reinoso, R. (2016). *Realidad Aumentada Posibilidades y usos en Educación*, CAVA.
11. Tralbaldo, S., Kamijo, M., & Fernández, A., (2015). *Mobile learning: nuevas realidades en el aula*, Editorial Oceano.

Tecnologías aplicadas a educación en UNNOBA

Claudia Russo^{1,2}, Tamara Ahmad¹, Mariana Ado¹, Paula Lencina¹, Eliana Serrano¹, Marina Rodríguez¹, Pedro Iglesias², Ana Smail¹, Natalia Bendatti², Nicolás Alonso¹, Rosana Piergalini¹, Trinidad Picco¹, Mercedes Guasch¹, Cecilia De Vitto¹, Laura Yamel², Pero Ochipinti¹, David Fernandez², Nicolás Moretti²

Instituto de Investigación y Transferencia de Tecnología (ITT)⁴
Comisión de Investigaciones Científicas (CIC)
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

Sarmiento y Newbery, Junín (B) – TE: (0236) 4477050

{claudia.russo, tamara.ahmad, mariana.ado, paula.lencina, eliana.serrano, marina.rodriguez, pedro.iglesias, ana.smail, natalia.bendatti, nicolas.alonso, rosana.piergalini, trinidad.picco, mercedes.guasch, cecilia.devitto, laurayamel, pedro.ochipinti, david.fernandez, nicolas.moretti}@itt.unnoba.edu.ar

Resumen

En este trabajo se muestran los avances realizados en las líneas de investigación correspondientes al proyecto “Informática y Tecnologías Emergentes”, de la Universidad Nacional del Noroeste de la provincia de Buenos Aires (UNNOBA). La tecnología aplicada a educación, tiene entre sus objetivos diseñar e implementar estrategias tecnológicas tendientes a minimizar la deserción en las carreras de la Universidad. En este contexto a partir de la utilización de la minería de datos se define un modelo de intervención áulica virtual que permite diseñar e implementar estrategias tecnológicas. Actualmente se trabaja en la utilización de avatares como tutores inteligentes en un entorno virtual 3D (EV3D), el uso del robot educativo programable realizado por la UNNOBA

(REP) y el desarrollo de píldoras educativas inclusivas. Estas experiencias se llevan a cabo en asignaturas de los primeros años de las carreras de informática.

Palabras clave: Tecnologías emergentes, Avatares, EV3D, píldoras educativas, REP, programación imperativa.

Contexto

En este documento se describen tareas ejecutadas en el marco de diferentes líneas de investigación dentro del proyecto denominado “Informática y Tecnologías Emergentes”. El mismo está acreditado y financiado por la UNNOBA en la convocatoria “Subsidios de Investigación Bianuales” 2017. En

¹ Docente Investigador en el Instituto de Investigación y Transferencia en Tecnología (ITT) / Escuela de Tecnología / UNNOBA

² Investigador Asociado Adjunto sin director a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

³ Becario PROMING / ITT / Escuela de Tecnología

⁴ Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

relación a lo anterior se desarrollan diferentes actividades que tienen lugar de

1. Introducción

Con el avance del tiempo se puede observar que la tecnología evoluciona a un ritmo que impacta directamente en la vida cotidiana de toda la sociedad. Lo hace abarcando diferentes ámbitos, impactando en la gran mayoría de las tareas y englobando todos los estratos sociales.

A partir de allí cobra verdadero sentido plantearse el concepto de Tecnologías Emergentes. A continuación se cita una definición que más se adecua a este trabajo: “Tecnologías emergentes son innovaciones en desarrollo que como su nombre lo dice en un futuro cambiarán la forma de vivir del ser humano brindándole mayor facilidad a la hora de realizar sus actividades, conforme la tecnología vaya cambiando estas también irán evolucionando logrando complementarse con la tecnología más moderna para brindar servicios que harán la vida del hombre mucho más segura y sencilla.”¹

En este sentido, el hecho de reciclar tecnología existente con el fin de darle un nuevo uso, implica obtener tecnología emergente para obtener mayores y nuevos beneficios.

Ante este contexto el impacto que este nuevo concepto acarrea en la sociedad es innegable. Pueden encontrarse muchas opiniones al respecto, dentro de un amplio abanico de fundamentaciones. Y también se advierten innegables consecuencias en el modelado de la sociedad, en las costumbres de los pueblos, en las diversas formas de relacionarse y comunicarse, la manera en

trabajo en el Instituto de Tecnología y Transferencia (ITT) de la UNNOBA.

la que las empresas producen, las metodologías de enseñanza y aprendizaje, entre otros aspectos.

Por otra parte el hecho de encontrar demandas laborales que buscan perfiles con buenas habilidades tecnológicas no es disparatado. La tecnología está en todos lados, en los Smartwatches, Smartphones, Tablets, heladeras y lavarropas de tipo “Smart”, TVs., etc.

Sin embargo, ante toda esta realidad, es necesario advertir que la forma de interactuar con la tecnología está inmersa en un cambio continuo. La manera en que se opera un ordenador va más allá de “sentarse en un escritorio frente a un monitor”, actualmente esa interacción implica un dinamismo en el que el usuario toma a la tecnología como una extensión de su cuerpo físico. Un ejemplo muy sencillo es pensar en qué grado se depende de las agendas electrónicas, de las notificaciones que le recuerdan al sujeto eventos, reuniones, cumpleaños, fechas que le permiten “descansar” en un sencillo artefacto electrónico su propia memoria.

Se relatan, en el siguiente apartado, diferentes trabajos que se realizan en la UNNOBA y están incluidos en un proyecto de I+D+i (Investigación, Desarrollo e innovación). El mismo nace de los resultados obtenidos en el proyecto “Tecnologías Exponenciales en contextos de realidades mixtas e interfaces avanzadas” desarrollado durante la convocatoria SIB 2015-2016.

Se proponen entonces objetivos que incluyen la identificación, contextualización, evaluación, desarrollo y aplicación de diversas herramientas informáticas en tecnologías emergentes. Estas tendrán un impacto directo en áreas tales como: telecomunicaciones, salud,

¹ Extraído del sitio “Reinventing Higher Education”, <https://www.ie.edu/IE/site/php/es/>

seguridad, gobierno, educación, industria, entre otras.

2. Líneas de Investigación, Desarrollo e Innovación

El proyecto de investigación "Informática y Tecnologías Emergentes", abarca cuatro líneas principales de investigación, desarrollo e innovación. Estas son: Tecnologías exponenciales (IT), Tratamiento masivo de datos (Big Data), Tecnología en educación (e-Tecnología), Robótica e Interacción Hombre-Máquina (HCI).

A continuación se enumeran las líneas que incluyen los trabajos aquí expuestos y se relata brevemente cada una de ellas:

Tecnología en educación (e-Tecnología): incluirá la conceptualización, análisis y desarrollo de software y metodologías. Comprenderá diferentes aspectos de la gestión del conocimiento en sistemas de e-learning, evaluación de calidad, el diseño de contenidos, diseño de juegos interactivos (gamificación), sistemas colaborativos y objetos de aprendizaje, metodologías y estrategias didácticas de manera de favorecer nuevos escenarios educativos.

Bajo esta línea se trabaja en el diseño de Avatares como tutores virtuales. Puntualmente se lo hace bajo el estudio de caso de la asignatura "Análisis y Diseño de Sistemas (ADS)" (materia de las carreras de informática en la UNNOBA). El objetivo de este trabajo es definir estrategias para la utilización de avatares en el desarrollo de actividades colaborativas dentro de un EV3D, y se hace bajo la indagación de la función pedagógica de los mismos en el contexto universitario y el rol tutorial que suponen. Además, se espera poder definir estrategias de incorporación de avatares en las carreras de Informática de la UNNOBA y diseñar actividades colaborativas a aplicar en el marco de

ADS. Con la puesta en marcha de una experiencia real con alumnos en dicha asignatura, se espera determinar el rol tutorial del docente virtual y poder analizar los intereses y aplicaciones de actividades colaborativas del EV3D en ADS.

Asimismo, bajo la misma línea de investigación se trabaja en el estudio de las píldoras educativas como contenido educativo digital potencialmente accesible. Se intentará responder si es posible definir una metodología para la generación de píldoras educativas inclusivas, es decir accesible para estudiantes con dificultades motrices, auditivas y/o visuales; y su aplicación en la asignatura Resolución de Problemas de Informática del Taller de Introducción y Articulación a los Estudios Superiores (TIAES) de la UNNOBA.

Por otra parte, la línea "Robótica e Interacción Hombre-Máquina (HCI)" incluye la investigación, diseño, construcción y programación de robots como solución a problemas multidisciplinares y transversales. Además, encierra la investigación de metodologías y herramientas que contribuyan a un cambio de paradigma hacia la interacción ubicua, entornos inteligentes y tecnologías interactivas de apoyo, prácticamente cualquier aspecto de la vida y las actividades humanas desde una perspectiva global y social.

Bajo esta última línea de investigación y ampliando el uso de la robótica en educación se trabaja en busca de estrategias educativas con respecto al paradigma imperativo de la programación de computadoras. Se lo hace utilizando como herramienta al REP. Particularmente se estudiará el caso de la asignatura "Introducción a la Programación Imperativa", perteneciente

a las carreras de informática de la UNNOBA.

Para ello se indaga acerca de ciertos temas que pueden enumerarse dentro de los siguientes puntos:

- Uso de robots como herramienta educativa.
- Tipos y niveles de interfaces, su uso en el plano educativo.
- Conceptos y habilidades para el desarrollo de interfaces de robots.
- Utilización de estrategias en propuestas educativas para promover habilidades en programación imperativa.
- Estrategias para la motivación de los estudiantes.

Puntualmente con este trabajo se espera que enriquezca las propuestas educativas para la enseñanza de la programación imperativa en la UNNOBA. Con tal fin se elaborarán estrategias educativas que utilicen al Robot Educativo Programable como herramienta y luego se las analizarán evaluando la posibilidad de su aplicación en las asignaturas “Introducción a la Programación Imperativa” y “Programación Imperativa” correspondiente al primer año de las carreras informáticas de la UNNOBA.

Sin embargo, este trabajo del uso del REP en la enseñanza de la programación de computadoras también está inmerso en la línea “e-Tecnología”, dada la naturaleza del objetivo principal al que apunta el trabajo: “Establecer estrategias educativas con respecto al paradigma imperativo de la programación de computadoras utilizando como herramienta al REP”.

3. Resultados obtenidos/esperados

Hasta el momento se desarrolló la conceptualización y puesta en marcha de la experiencia en el EV3D para luego poder definir las estrategias pedagógicas

del rol tutorial de los avatares. Con respecto al uso de píldoras educativas, se logró evaluar las metodologías existentes y formular una metodología para la generación de píldoras educativas inclusivas. En cuanto al REP, también se realizó la conceptualización de la implicancia del uso del REP en las asignaturas de programación y se está desarrollando la interfaz que permite la conexión del REP con el lenguaje de programación utilizado.

4. Formación de Recursos Humanos

El equipo de trabajo está compuesto por investigadores formados y en formación, becario de posgrado y becarios de grado. En relación a las líneas de investigación presentadas se espera que en el transcurso de este año se concreten tres tesis de la “Maestría de Tecnología Aplicada a la Educación” de la Universidad Nacional de La Plata (UNLP), una tesis de la “Licenciatura en Sistemas” de la UNNOBA, se contribuya al inicio de una tesis de la “Maestría en Educación en Entornos Virtuales” de la Universidad Nacional de la Patagonia Austral y la culminación de una tesis doctoral perteneciente al Doctorado de ciencias informáticas de la UNLP.

5. Bibliografía

BARROSO JEREZ, Clara; Criterios pedagógicos en el uso de multimedia en educación: los agentes pedagógicos; IV Ponencia: Los lenguajes de las pantallas. Impacto en las relaciones sociales de los jóvenes y retos educativos; XXII Seminario interuniversitario de teoría de la educación. Otros lenguajes en educación; Sitges; Universitat de Barcelona; Barcelona; 2003.

CARRERA, F. (Dir.). Knowledge pills: Metodología de píldoras de conocimiento (Lifelong Learning Programme's Report).

European Commission: Fundacja Obserwatorium Zarządzania; Prestin – Preparação de Estudos e Investimentos, Lda.; ZEUS Consulting S.A.; Instituto Tecnológico de Aragón; y Nowoczesna Firma S.A. 2011.

GARCÍA ARETIO, Lorenzo, RUIZ CORBELLA, Marta, GARCÍA BLANCO, Miriam; Claves para la educación: actores, agentes y escenarios en la sociedad actual; Narcea SA y Universidad Nacional de Educación a Distancia coeditores; Madrid; 2009.

MAESTRI, George; Creación digital de personajes animados: técnicas avanzadas; Anaya Multimedia; Madrid; 2002.

POVEDA CRIADO, M Y THOUS TUSET, M. Mundos virtuales y avatares como nuevas formas educativas. Historia y Comunicación Socia. Vol. 18. N° Esp. Nov. (2013) 469-479. ISSN: 1137-0734. Recuperado de: http://dx.doi.org/10.5209/rev_HICS.2013.v18.44262; 2013

PRENSKY, Marc; Nativos Digitales, Inmigrantes Digitales; MCB UniversityPress, Vol. 9 No. 6.; Madrid; 2001.

RUSSO, C.; SAROBE, M.; AHMAD, T.; LOMBARDO, M.; LENCINA, P.; RAMÓN, H. Experiencia del uso del EV3D en UNNOBA. TE&ET. Argentina. 2016.

RUSSO, C., RAMÓN, H., LENCINA, P., ADÓ, M., IGLESIAS, P. Programando con el robot educativo en la UNNOBA, TE &EY Argentina, 2017.

SERRANO, E., SAROBE, M., RUSSO, C. C., TESSORE, J. P., & RAMÓN, H. D. La Inclusión como desafío en el inicio de la vida universitaria. In X Congreso sobre Tecnología en Educación & Educación en Tecnología (TE & ET). Corrientes, 2015.

REALIDAD AUMENTADA Y OTRAS TECNOLOGÍAS EMERGENTES EN PROCESOS DE ENSEÑANZA Y APRENDIZAJE. APROXIMACIONES METODOLÓGICAS AL DISEÑO Y EVALUACIÓN DE PROPUESTAS DIDÁCTICAS

Cecilia Sanz²; Tatiana Gibelli^{1,3}; Edith Lovos¹; Paula Suárez¹; Álvaro Saldivia;
Sergio Condó; Verónica Cuevas³

¹Centro Interdisciplinario en Derechos, Inclusión y Sociedad (CIEDIS)
Sede Atlántica, Universidad Nacional de Río Negro,

² Investigador asociado de la Comisión de Investigaciones Científicas de la
Provincia de Buenos Aires. Instituto de Investigación en Informática LIDI (III-
LIDI), Facultad de Informática, Universidad Nacional de la Plata

³ Centro Universitario Zona Atlántica (CURZA)
Universidad Nacional del Comahue

csanz@lidi.info.unlp.edu.ar - {tgibelli, elovos, psuarez, scondo, asaldivia}@unrn.edu.ar -
vcuevas1976@gmail.com

RESUMEN

La implementación de propuestas didácticas que incluyen tecnologías innovadoras requieren de un cuidadoso y adecuado diseño, seguimiento y evaluación del proceso. Desde un proyecto de investigación acreditado por la Universidad Nacional de Río Negro (UNRN) se propone hacer un aporte en este sentido. Este proyecto, a través de una metodología de investigación-acción, analiza y busca conocer el impacto de diferentes tecnologías digitales en contextos educativos específicos. Se pone el énfasis en tecnologías emergentes como la realidad aumentada, explorando metodologías de inclusión, diseño de propuestas y evaluación, analizando sus posibilidades y barreras en procesos educativos. En este trabajo se presentan avances del proyecto, resultados preliminares y líneas de trabajo futuro.

Palabras clave: Tecnologías de la Información y la Comunicación, Realidad Aumentada, Educación, Metodologías

CONTEXTO

Esta investigación se desarrolla en el marco del proyecto bianual (2016-2018) “La mediación de las tecnologías de la información y la comunicación en procesos educativos. Innovaciones para mejorar los procesos de enseñanza y aprendizaje”, acreditado y financiado por la UNRN.

1. INTRODUCCIÓN

Las TIC han provocado un alto impacto en el sistema educativo. Actualmente nos encontramos con un perfil de alumno cuyo desarrollo personal, social y laboral ha estado –y estará- fuertemente mediado por las TIC,

lo que influye, sin lugar a dudas, en el modo en que entiende y analiza el mundo que le rodea. Y a diferencia de otros tipos de tecnologías, como sostiene Silva (2013), éstas se caracterizan por permear el comportamiento de todos sus actores (estudiantes y docentes) en su vida cotidiana. En este punto, desde la perspectiva docente, las TIC además de brindar la posibilidad de realizar ciertas actividades de manera más rápida, cómoda o fiable, nos permiten crear nuevas estrategias didácticas y, con ellas nuevos entornos de aprendizaje, y es aquí donde reside su verdadero potencial transformador (Cabero Almenara, 2010).

Sin embargo no hay que caer en una visión tecnocrática, tal como destaca Sánchez (2002), la integración curricular de las TIC implica tener una filosofía subyacente, un proyecto educativo que implique un proceso de cambio e innovación educativa, un aprender de contenidos específicos, modelos de aprender para lograr "la invisibilidad de la tecnología para una visibilidad del aprender.". Y como sostiene Silva (2013), en relación a la inclusión de las TIC en educación el foco de estudio debe ponerse sobre las metodologías y los estudiantes.

Una de las tecnologías emergentes con posibilidad de inclusión en Educación es la Realidad Aumentada. Según Azuma et al. (2001), la RA se caracteriza por: (a) una combinación de objetos virtuales y reales en un escenario real, (b) usuarios interactuando en tiempo real y (c) una alineación entre los objetos reales y virtuales. Respecto a su potencialidad en el plano educativo, la RA permite incorporar multimedia a los procesos de enseñar y aprender, innovar en la práctica docente y promover el diseño de materiales educativos atendiendo a los requerimientos didácticos (Avendaño, 2012). Cabero y Barroso (2015), señalan que para su incorporación en los procesos

de enseñanza y aprendizaje es necesario tener en cuenta que el diseño de los entornos debe ser flexible de manera que su inclusión en el espacio de enseñanza y aprendizaje, no se convierta en un problema técnico.

La RA puede proporcionar grandes oportunidades en distintas áreas como la ciencia o la ingeniería, puesto que estas disciplinas conllevan en su currículum un enfoque práctico en el aula (Andújar et al., 2011). Sin embargo, como señala Prendes Espinosa (2015), aun siendo la tecnología lo suficientemente madura para ser útil en los entornos educativos, los nuevos desarrollos de RA llevarán a una nueva generación de aplicaciones que desplegarán mucho más su potencial.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El objetivo principal del proyecto es investigar metodologías y estrategias innovadoras que favorezcan procesos educativos mediados por TIC.

Los objetivos específicos son:

- Explorar marcos teóricos que permitan interpretar contextos educativos mediados por TIC.
- Evaluar la mediación de TIC en situaciones educativas concretas.
- Proponer innovaciones tecnológicas y metodológicas para necesidades educativas específicas, en particular, relacionadas con realidad aumentada y uso de dispositivos móviles.
- Promover el desarrollo de prácticas docentes innovadoras con apoyo de recursos informáticos.

Se espera que las metodologías y estrategias investigadas permitan potenciar los procesos de enseñanza y aprendizaje y el desarrollo de habilidades

(pensamiento crítico, trabajo en equipo, comunicación, etc) necesarias para desarrollarse en la sociedad del conocimiento.

3. RESULTADOS

Durante el primer año del proyecto se llevó adelante una revisión bibliográfica sobre el tema por parte de los integrantes del proyecto. Los temas abordados en esta revisión fueron las tecnologías emergentes, en particular la realidad aumentada, y su uso en diferentes contextos educativos. La revisión también consideró las áreas de aplicación específicas acorde a la formación y espacios de desempeño de los docentes participantes en el proyecto, por esto se focalizó especialmente en las áreas de matemática y programación.

Luego se exploraron aplicativos y herramientas que se pueden utilizar para el diseño de propuestas didácticas con realidad aumentada. En este sentido se llevaron adelante las siguientes actividades investigativas:

1. Análisis comparativo entre dos herramientas que permiten el diseño de actividades con RA por geolocalización usando dispositivos móviles posibilitando el diseño de juegos y actividades en las que existe un recorrido y toma relevancia el posicionamiento del jugador. Los resultados del mismo han presentados en el Congreso Iberoamericano de Educación y Sociedad (CIEDUC 2017).
2. Exploración de herramientas de software disponibles para realizar modelos 3D que se insertarán en actividades de RA teniendo en cuenta la necesidad de diseñar los propios modelos. Se exploraron herramientas de software con énfasis

en aquellas que permiten el trabajo con funciones matemáticas en tres dimensiones, profundizando en dos de ellas: Blender y SketchUp. Los resultados se presentaron en el XXIII Congreso Argentino de Ciencias de la Computación (Gibelli, Graziani, y Sanz, 2017).

3. Revisión de materiales educativos digitales (MED) basados en RA que puedan utilizarse como apoyo al aprendizaje de conceptos básicos de programación en el contexto universitario. Los MED explorados se pusieron a prueba en dos experiencias de aprendizaje sobre estructuras de control y manejo de estructuras de datos dinámicas (listas). Los resultados fueron aceptados para su presentación al Sexto Congreso Argentino de la Interacción Persona Computador@, telecomunicaciones, Informática e Información Científica (IPCTIIC 2017)
4. Revisión de instrumentos y variables para evaluar materiales educativos digitales que incorporan realidad aumentada (Lovos, Gibelli y Sanz, 2017). En base a esta revisión, actualmente se está en etapa de realización de diseño de un instrumento propio de recolección de datos para poder utilizar en la evaluación de las futuras implementaciones.

Por otra parte, se comenzó con el diseño de propuestas y materiales educativos con uso de realidad aumentada. En este sentido podemos mencionar el diseño de dos propuestas con uso de RA. Una de ellas es una actividad para el aprendizaje de conceptos de geometría tridimensional como puntos, vectores, planos y rectas en el espacio. La actividad se basa en establecer asociaciones entre objetos cotidianos con conceptos de geometría

tridimensional. Para facilitar estas asociaciones se hizo uso del aplicativo Aurasma¹ que permitió generar “Auras” a través de los cuales se agrega información virtual a los objetos físicos considerados. La otra propuesta consiste en un material digital para la enseñanza de temas de cálculo, específicamente, la resolución de problemas de optimización. Se trata de una guía de estudio, que incluye varios problemas resueltos y otros tantos propuestos, en orden de dificultad creciente. Cada problema resuelto contiene un código QR que enlaza con un sitio web de Geogebra donde se accede a una animación en la cual se puede apreciar claramente lo que plantea el problema, y encontrar el resultado con facilidad. Ambas actividades están destinadas a estudiantes universitarios y se realizaron experiencias pilotos durante el segundo cuatrimestre de 2017 con estudiantes de la Sede Atlántica de la UNRN, en particular de las carreras de Contador Público, Licenciatura en Sistemas e Ingeniería Agronómica. En base a dichas experiencias con adecuaciones al diseño, se prevé realizar nuevas implementaciones durante el primer cuatrimestre de 2018.

Actualmente se están llevando adelante diversas implementaciones de los diseños realizados y recolectando datos de valoración por parte de los estudiantes y docentes involucrados para su posterior análisis. Como trabajo futuro se prevé la realización de informes técnicos y publicaciones que den cuenta de las metodologías diseñadas, sus fortalezas y debilidades. Consideramos que esto es un aporte en la temática y puede resultar de interés para investigadores y docentes que estén trabajando en temas de realidad

umentada aplicada en contextos educativos.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto de investigación es dirigido por la Dra Cecilia Sanz, docente de la Facultad de Informática de la Universidad Nacional de la Plata. El equipo de investigadores está formado por: dos docentes con dedicación completa y tres docentes concurrentes de la UNRN, una docente externa a la institución y dos alumnos de la carrera Lic. en Sistemas de la UNRN.

En cuanto a la formación profesional de los integrantes, la dirección externa con experiencia y antecedentes relevantes en la temática del proyecto constituye un aporte importante a la formación del equipo local. En este sentido, durante el primer cuatrimestre de 2017 la directora dictó el curso el postgrado “Introducción a la Realidad Aumentada. Posibilidades para escenarios educativos” en la Sede Atlántica de la UNRN. El mismo contó con 12 inscriptos de los cuales 10 completaron y aprobaron todas las actividades previstas. La experiencia del curso ha resultado satisfactoria y ha permitido la profundización de la investigación en estas temáticas para los participantes y en especial para los integrantes del proyecto. Asimismo se han realizado cursos de postgrado del Doctorado en Ciencias Informáticas (UNLP) y otros vinculados a la temática del proyecto, sus intereses personales y formación específica.

Además, se realizan aportes a la formación de nuevos investigadores. En particular se participa en la dirección y asesoramiento de diversas tesis de posgrado de la Maestría en Tecnología Informática Aplicada en Educación.

¹<https://www.aurasma.com/>

Asimismo se realiza asesoramiento a alumnos de grado a través de dos becas de inicio a la investigación UNRN y una beca CIN convocatoria 2017 aprobada.

5. REFERENCIAS

- Andújar, J. M.; Mejías, A.; Márquez, M. A. (2011). Augmented Reality for the Improvement of Remote Laboratories: An Augmented Remote Laboratory. *IEEE Transactions on Education*, 54 (3), (492-500)
- Avendaño, V. y Domínguez, L. A. (2012). Realidad aumentada: Una exploración al escenario de la virtualidad educativa. Madrid: Editorial Académica Española.
- Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *Computer Graphics and Applications*, IEEE, 21(6), 34-47.
- Cabero Almenara, J. (2010). Los retos de la integración de las TICs en los procesos educativos. Límites y posibilidades. *Perspectiva educacional*, 49(1), 32-61.
- Cabero, J. y Barroso, J. (2015). Realidad Aumentada: posibilidades educativas. En Ruiz-Palmero, J., Sánchez-Rodríguez, J. y Sánchez-Rivas, E. (Edit.). *Innovaciones con tecnologías emergentes*. Málaga: Universidad de Málaga.
- Gibelli, T. I., Graziani, A., & Sanz, C. V. (2017). Revisión de herramientas para la creación de modelos 3D orientados a la enseñanza de la matemática con realidad aumentada. In XXIII Congreso Argentino de Ciencias de la Computación (CACIC, La Plata, 2017).
- Lovos, E., Gibelli, T. I., & Sanz, C. V. (2017). Evaluación de materiales educativos digitales que incorporan realidad aumentada: revisión de variables e instrumentos. In XII Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET, La Matanza 2017).
- Sánchez, J. (2002). Integración curricular de las TICs: conceptos e ideas. *Actas VI Congreso Iberoamericano de Informática Educativa, RIBIE* (pp. 20-22).
- Silva, J. (2013). TIC en educación superior. Una reflexión teórica-práctica. En *Revista de la Asociación Colombiana de Ciencias(Col.)*, 25: 92-97.
- Prendes Espinosa, C. (2015). Realidad aumentada y educación: análisis de experiencias prácticas. *Pixel-Bit. Revista de Medios y Educación*, (46), 187-203.

RECURSOS TIC PARA EL APRENDIZAJE DE LA QUIMICA Y LA FISICA EN EL CICLO BÁSICO UNIVERSITARIO

María I. Vera¹, Irene Lucero²; Marta G. Stoppello²; Raquel H. Petris³ y Liliana I. Giménez¹

¹Departamento de Química. Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste. Argentina.

² Departamento de Física. Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste. Argentina.

³ Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste. Argentina.

marile.vera5@gmail.com; irmairene2005@yahoo.com.ar; mstopello@hotmail.com;
raquelpetris@hotmail.com; bioliligi@yahoo.com.ar

RESUMEN

El Proyecto 17F001/14, SGCyT (UNNE): “*Innovación con TIC para fortalecer la enseñanza y aprendizaje de las actividades prácticas de Química y de Física en los primeros años de FaCENA*” surge como un nuevo desafío del grupo de investigación para consolidar diferentes formas de acceso al conocimiento con el uso de TIC; en esta ocasión, con propuestas que plantean distintas concepciones sobre la enseñanza y aprendizaje de los trabajos prácticos. El propósito general es el de indagar acerca de la apropiación de conceptos fundamentales de distintos temas a través de las clases prácticas de problemas y/o experimentales con uso de recursos TIC en asignaturas de Química y Física de los primeros años de carreras de grado de la FaCENA, y su posterior evaluación en relación con los rendimientos académicos.

Entendemos que es una buena ocasión para introducir procesos de innovación con TIC en el caso de clases masivas de alumnos ingresantes. Pero la mera introducción de TIC no cambiará el aprendizaje de forma significativa si no posee un proyecto intencional y deliberado de cambio, si no modificamos nuestras prácticas pedagógicas.

Palabras Clave: *Recursos TIC; enseñanza; aprendizaje; ciencias experimentales; innovación.*

CONTEXTO

Las investigaciones se realizan en el marco del Proyecto PI: 17F001/14, “*Innovación con TIC para fortalecer la enseñanza y aprendizaje de las actividades prácticas de Química y de Física en el Ciclo Básico de FaCENA*”. Fue aprobado por la Secretaria General de Ciencia y Técnica (SGCyT) de la Universidad Nacional del Nordeste (UNNE), Resolución N° 358/15 C.S. Período 2015-2018.

Se indaga sobre la aplicación de Recursos TIC y su efecto en el aprendizaje de las Asignaturas Química General, correspondiente a primer año de carreras de Ingeniería (Eléctrica, en Electrónica y en Agrimensura); Laboratorio de Calor y Termodinámica de alumnos que cursan las carreras de Profesorado y Licenciatura en Física y Física Atómica para alumnos de Ingeniería en la Facultad de Ciencias Exactas y Naturales y Agrimensura (FaCENA) de la UNNE.

INTRODUCCIÓN

Enseñar ciencias nunca ha sido una tarea fácil, pero parece que los retos se multiplican en estos tiempos de cambios acelerados, tanto en lo referente a los conocimientos que hay que enseñar o en los mejores métodos para hacerlo como en lo que respecta al alumnado a quien se dirige la enseñanza, e incluso en las demandas que la sociedad plantea (Jiménez Aleixandre, 2003).

En Argentina, con las dos reformas educativas -la Ley Federal de Educación y la actual Ley de Educación Nacional- se ha impuesto en la educación secundaria el concepto de alfabetización científica y el enfoque ciencia tecnología sociedad, para la enseñanza de las ciencias naturales. Esta cuestión ha disminuido “la importancia de los contenidos tradicionalmente considerados como estrictamente disciplinares, para dar espacio curricular a aspectos situados más en el campo de la comprensión pública de la ciencia” (Caamaño, 2005). Tal es así que el joven ingresante a la universidad no está acostumbrado al pensamiento abstracto, al manejo de simbología y ecuaciones algebraicas con significado conceptual tanto en el campo de la Química como de la Física. Como la tecnología forma parte de la cotidianidad de los individuos, debe ser tenida en cuenta en los procesos formativos; de lo contrario el estudiante se cuestionará si le sirve de algo lo que está aprendiendo o si lo que está viviendo en las aulas tiene relación con su quehacer, dejándole grandes vacíos que involucran desatención y desmotivación (Cruz y Espinosa, 2012). Los estudiantes pueden acceder a través de internet a una enorme cantidad de información pero esta información no equivale a conocimiento, ya que para que esta se convierta en conocimiento necesita de un profesor que guíe al estudiante en este proceso (Cebrián, 2003). Según Alejandro Alfonso (2004, en Cruz y Espinosa, 2012) no se puede involucrar únicamente tecnología, pues por sí sola no logra los objetivos esperados; sin un enfoque pedagógico adecuado, estas mismas tecnologías podrían tener un efecto negativo. Las posibilidades que las TIC ofrecen para la enseñanza y la formación en el terreno de la química y la física son diversas, y van desde facilitar la comunicación profesor-estudiante, hasta presentar información o desarrollar entornos específicos como pueden ser los laboratorios virtuales (Cabero, 2007). La química y la física son disciplinas que forman parte del diseño curricular de un gran número de carreras universitarias. Con el uso de las computadoras han aparecido nuevas

formas de aprendizaje para la enseñanza de ambas disciplinas que posibilitan el acercamiento y motivación de los estudiantes por una tarea autodidacta para el aprendizaje de las ciencias. Si bien el uso del laboratorio en la enseñanza de la química resulta indispensable, coincidimos con Cabero (2007) quien reconoce algunas dificultades, especialmente, en la enseñanza presencial de nivel universitario inicial debido -entre otras causas- a: a) El número de estudiantes por cada grupo con clases numerosas y recursos edilicios y humanos insuficientes, b) Los recursos económicos disponibles para la compra de reactivos necesarios, c) la heterogeneidad de los estudiantes en cuanto a perfiles de las carreras que comparten el cursado. Surge así un nuevo paradigma de trabajo virtual que se complementa con las clases y la tarea de aula. Esta nueva forma de interacción tiene tres componentes básicos para la herramienta software: los laboratorios virtuales, los programas de modelación y los simuladores que apoyan los procesos de enseñanza y facilitan la tarea al docente (Cataldi, 2009). El laboratorio virtual o el simulador son verdaderos ambientes protegidos que evitan los riesgos propios de exposición física debido a la manipulación de materiales y reactivos, permitiendo que los estudiantes ensayen, prueben y se arriesguen a equivocarse sin miedos. Respecto de la Física, el uso del experimento, se constituye en herramienta esencial para la enseñanza y comprensión de la Física ya que proporciona al estudiante un pensamiento más creativo y confianza por la investigación científica, lo que le permitirá descubrir y comprobar determinados fenómenos o principios científicos. En los experimentos presenciales -los que se llevan a cabo en el laboratorio- el estudiante tiene la posibilidad de interactuar con los elementos de medición y estudiar el fenómeno mediante la interacción práctica; los experimentos virtuales responden a un tipo de prácticas que actualmente está en auge gracias al desarrollo de las TIC; en general son programas de computadora que brindan alternativas al docente para mostrar y enseñar un fenómeno natural mediante la

visualización de los diferentes estados que el mismo puede presentar (Ubaque Brito, 2009, en Cruz y Espinosa, 2012). Este tipo de herramientas de experimentación les permiten a los estudiantes cambiar las condiciones de un problema determinado sin poner en riesgo su integridad ni la de los equipos técnicos que intervienen.

Disponer de una computadora conectada a internet es expandir las paredes del aula, pudiendo producir y consumir contenidos dentro y fuera de ella. Aparecen así nuevas formas de presentar contenidos: textos digitales, hipertextos, infografías, videos.

El video es un recurso didáctico que combina imágenes y sonidos permitiendo visualizar procesos o procedimientos. Según Marqués (1999), se denomina video educativo a los materiales videográficos que pueden tener una utilidad en educación, incluyendo en este concepto a los videos didácticos y cualquier otro tipo, que pueda resultar útil a la enseñanza. Dentro de la clasificación de videos que presenta este autor, la lección monoconceptual y la lección temática son formas de video muy útiles, dado que no presentan larga duración y se refieren a un tema específico que es presentado en forma sistemática y con la profundidad adecuada a los destinatarios. Con esta filosofía *“es posible elaborar videos propios con fines didácticos, donde el profesor dé explicaciones detalladas de manera corta y simple de un determinado tema, presentando por ejemplo, la forma de resolución de problemas específicos”* (Vera et al, 2016).

Un video didáctico debe estar pensado para usarse en conjunto con otros materiales didácticos (Cebrian, 1994) que manejen la misma información. En la postura que aquí se asume, el video es complemento de los textos recomendados, de las guías de problemas, del discurso en las clases presenciales, buscando unificar y clarificar nomenclaturas y procedimientos de cálculo.

Una forma de uso de videos educativos se da en la flipped classroom o aula invertida (Bergmann, J. y Sams, A., 2012). El aula invertida es *“un modelo pedagógico que transforma ciertos procesos que de forma*

habitual estaban vinculados exclusivamente al aula, transfiriéndolos al contexto extraescolar. Es decir, invierte la forma tradicional de entender una clase: aquellas actividades ligadas principalmente a la exposición y explicación de contenidos pasan a ofrecerse fuera del aula, por medio de herramientas tecnológicas como puede ser el video” (García Barrera, 2013)

Según Velásquez Huerta (2010) la incorporación de las TIC, como cualquier otro material, debe ser necesariamente evaluada y considerada dentro de la programación curricular, teniendo en consideración que el desarrollo de los contenidos posibilite su empleo en las actividades de aprendizaje.

La utilización de una mayor diversidad de recursos materiales, tecnologías y métodos de enseñanza (páginas web interactivas con contenidos didácticos, tutorías, seguimiento del alumno, etc.), ha permitido mejorar la calidad de resolver problemas y tomar decisiones, de aprender y trabajar en equipo, de aplicar conceptos en la práctica, y en definitiva mejorar el proceso enseñanza-aprendizaje (Guzman Parra, 2009).

El grupo de Investigación implementa desde el año 2012, el uso del Aula Virtual Ecaths en asignaturas de primer año, para favorecer la comunicación docente – alumno y el proceso de enseñanza –aprendizaje. Ecaths (<http://www.ecaths.com/home.php>) - servicio de uso libre y gratuito- es un sistema de gestión online de Asignaturas cuya función principal es complementar la cursada presencial con un espacio virtual de interacción y construcción de conocimiento colectivo.

LINEAS DE INVESTIGACION

- Enseñanza y aprendizaje de la Química y la Física con recursos TIC en el ciclo básico universitario.
- Diseño de recursos didácticos con TIC, uso en las clases y evaluación del impacto del uso para el aprendizaje de contenidos de Química General.

- Se centra en el uso de videos explicativos monotemáticos, el uso de simulaciones para laboratorios virtuales, planillas de cálculo colaborativas para el tratamiento de datos en los laboratorios de Física. Diseño, elaboración y uso de Presentaciones Digitales en la instancia evaluativa de alumnos.

Nos planteamos llevar a cabo la investigación teniendo como base los siguientes principios generales:

- El aprendizaje debe ser activo, el alumno construirá su conocimiento a través de la colaboración, la interacción, la búsqueda de información y el contraste de puntos de vista.
- La tecnología siempre es un medio y no un fin en sí misma: lo importante no es el desarrollo de la técnica por la técnica, sino el uso que vamos a hacer de ella para mejorar los procesos de enseñanza aprendizaje.
- El papel del docente es el de un facilitador del aprendizaje, en el sentido de crear las condiciones oportunas, dar orientaciones, solucionar dudas o problemas, colaborando en la construcción del conocimiento compartido.

La evaluación del éxito de la innovación se basa en diferentes parámetros: la regularización y aprobación de las asignaturas involucradas, resultados académicos obtenidos con metodologías activas en comparación con la tradicional, entrevistas y encuestas de valoración que miden el grado de conformidad o satisfacción del alumnado con la innovación implementada.

RESULTADOS

Se analiza el impacto en el aprendizaje de los estudiantes del uso de videos educativos diseñados como recurso didáctico de apoyo a las clases presenciales. El análisis se hace en forma particularizada para cada uno de los temas que son evaluados en el primer parcial: a) Formulación y Nomenclatura química; b) Ecuaciones químicas y c) Estequiometría.

Los resultados han sido presentados en forma de ponencias y/o papers en los siguientes eventos científicos:

Año 2015: TE y ET 2015 (FaCENA, UNNE); Jornadas de intercambio de experiencias: las TIC en la educación universitaria (UNNE); XVII REQ (UnCAus); X Jornadas Nacionales y VII Jornadas Internacionales de Enseñanza de la Química Universitaria, Superior, Secundaria y Técnica (AQA, Buenos Aires).

Año 2016: III Congreso Argentino de Ingeniería – IX Congreso de Enseñanza de la Ingeniería (CADI, Facultad de Ingeniería, UNNE); SIEC 2016 (Vigo, España); Virtual USATIC (Zaragoza, España)

Año 2017: Primer Congreso Latinoamericano de Ingeniería (CONFEDI, UNER); Primer Congreso de Educación y Tecnologías del Mercosur. De la digitalización a la virtualización (Programa UNNE Virtual); XI Jornadas Nacionales y VIII Jornadas Internacionales de Enseñanza de la Química Universitaria, Superior, Secundaria y Técnica (AQA, Buenos Aires); 6° Congreso Internacional sobre Buenas Prácticas con TIC (modalidad virtual) (Universidad de Málaga, España).

En **2018** se tiene previsto participar en congresos nacionales e internacionales.

En forma general en los trabajos presentados se comprobó que los resultados de las evaluaciones son mejores que los obtenidos en instancias previas a la implementación de cada una de las innovaciones.

LINEAS FUTURAS: Se continuará con la elaboración de videos explicativos para temas que tradicionalmente son difíciles de comprender por parte del alumnado y se impulsará la aplicación de metodología flipped classroom en aquellos contenidos factibles de ser aplicada.

FORMACION DE RECURSOS HUMANOS

El equipo esta integrado por docentes investigadores de los Departamentos Pedagógicos de Química, Física e Informática perfiles que aportan fortaleza para las investigaciones complementándose.

Se cuenta con alumnos adscriptos que son iniciados en tareas inherentes a la Investigación Educativa. Se ha dirigido

diferentes tesis de grado finalizadas con resultados positivos.

BIBLIOGRAFIA

Bergmann, J. y Sams, A. (2012). *Flip Your Classroom: Talk To Every Student In Every Class Every Day*. Washington, DC: ISTE, citado en García Barrera, Alba (2013). El aula inversa: cambiando la respuesta a las necesidades de los estudiantes. Avances en Supervisión Educativa. Revista de la Asociación de Inspectores de Educación de España. N° 19. Disponible en: http://www.adide.org/revista/images/stories/revista19/ase19_mono02.pdf. Consulta agosto 2015

Caamaño, A. (2003). Los trabajos prácticos en ciencias. En Enseñar ciencias. Jiménez Aleixandre (coord), Caamaño, A.; Oñorbe, A.; Pedrinaci, E.. y de Pro, E. Barcelona: Grao.

Caamaño, A. (2005), en Galagovsky, L. (2007). Enseñar Química vs. Aprender Química: una ecuación que no está balanceada. Revista Química Viva. Vol 6, número especial. UBA. Disponible en: <http://www.redalyc.org/articulo.oa?id=86309909>. Consulta: Junio 2016

Cabero Almenara, J. (2007). Las TIC's en la enseñanza de la química: aportaciones desde la Tecnología Educativa. En Bodalo, A. y otros (Eds) 2007. Química, Vida y Progreso. Murcia: Asociación de Químicos de Murcia.

Cataldi, Z, Donnamaría M.C., Lage, F. (2009). Didáctica de la química y TICs: Laboratorios virtuales, modelos y simulaciones como agentes de motivación y de cambio conceptual. TEyET 2009. Facultad de informática UNLP. Disponible en: <http://sedici.unlp.edu.ar/bitstream/handle/10915/18979/11.pdf?sequence=1> Consulta: 2 de octubre de 2013

Cebrián, M. (Coord.)(2003). Enseñanza virtual para la innovación universitaria. Madrid: Narcea.

Cebrian, M. (1994). *Los videos didácticos: claves para su producción y evaluación*. Pixel-bit. N°1. Disponible en <http://www.sav.us.es/pixelbit/pixelbit/articulos/n1/n1art/art13.htm>. Consulta marzo 2016.

Cruz Ardila, J.C. y Espinosa Arroyave, V. (2012). Reflexiones sobre la didáctica en física desde los laboratorios y el uso de las TIC. Revista Virtual Universidad Católica del Norte, 35, 105-127. Medellín. Colombia.

Disponible en:

<http://www.redalyc.org/articulo.oa?id=194224362007>

García Barrera, A. (2013). El aula inversa: cambiando la respuesta a las necesidades de los estudiantes. Avances en Supervisión Educativa. Revista de la Asociación de Inspectores de Educación de España. N° 19. Disponible en:

http://www.adide.org/revista/images/stories/revista19/ase19_mono02.pdf. Consulta agosto 2015.

Guzmán Parra, V. (2009). Evolución del modelo docente: efectos de la incorporación del uso de una plataforma virtual, videos educativos y CD interactivos. Edutec Revista Electrónica de Tecnología Educativa. 30. <http://www.edutec.rediris.es/revelec2/revelec30/> Consulta: 1° de julio 2010.

Jiménez Aleixandre, M.P. (2003). Aprender a pensar científicamente. En Enseñar Ciencias. Jiménez Aleixandre (coord), Caamaño, A.; Oñorbe, A.; Pedrinaci, E.. y de Pro, E. Barcelona: Grao.

Marqués Graells, P. (1999). Los videos educativos: tipologías, funciones, orientaciones para su uso. Departamento de Pedagogía Aplicada. Facultad de Educación UAB. Disponible en:

<http://www.peremarques.net/videoori.htm>.

Consulta: Septiembre 2014

Velásquez Huerta, R.A. (2010). Las TICs en la educación universitaria. Material de estudio de la Diplomatura en Docencia Universitaria, 2da Edición. Universidad Nacional Daniel A. Carrión.

APLICACIÓN DE TÉCNICAS ESTADÍSTICAS Y DE MINERÍA DE DATOS PARA EL ANÁLISIS DE PERFILES DE RENDIMIENTO ACADÉMICO: EL CASO DE LA FACULTAD DE CIENCIAS EXACTAS Y NATURALES DE LA UNIVERSIDAD NACIONAL DE LA PAMPA

Lorena Verónica CAVERO⁽¹⁾, María Paula DIESER⁽¹⁾, María Cristina MARTÍN⁽¹⁾⁽²⁾, Sofía FUNKNER⁽¹⁾, Janina ROLDAN⁽¹⁾, Erica SCHLAPS⁽³⁾, Diamela TITIONIK⁽¹⁾, Laura WAGNER⁽¹⁾

⁽¹⁾ Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa

⁽²⁾ Departamento de Matemática, Universidad Nacional del Sur

⁽³⁾ Instituto de Ciencias Polares, Ambiente y Recursos Naturales, Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur

{cavero, pauladieser, maritamartin}@exactas.unlpam.edu.ar

RESUMEN

En el proceso de inscripción a las carreras de grado de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa, y a lo largo del recorrido académico que hacen los estudiantes por la institución, se recolectan múltiples datos a través de los sistemas de gestión. Éstos constituyen una importante fuente de información, en tanto se extraiga conocimiento útil para el análisis de la realidad de los estudiantes y los contextos en los que ellos aprenden, y para el diseño de eventuales planes de acción. En particular, interesa a la comunidad institucional la detección temprana de estudiantes en situación de riesgo en términos de deserción o retraso en el alcance del grado.

La línea de investigación presentada, propone estudiar y aplicar distintos métodos que ofrece la Minería de Datos, el Análisis de Datos Multivariados, la Teoría de Respuesta al Ítem y el Análisis de Supervivencia, sobre los datos registrados en el sistema de gestión de información estudiantil de la Institución con el propósito de caracterizar la trayectoria académica de los estudiantes, y detectar patrones compatibles con situaciones de dificultades en el aprendizaje, que puedan derivar en abandono de los estudios.

Palabras clave: minería de datos, análisis de datos multivariados, teoría de respuesta al ítem, análisis de supervivencia, deserción universitaria.

CONTEXTO

Durante el periodo 2014 - 2017 se realizaron tareas de investigación, en el ámbito de la Facultad de Ciencias Exactas y Naturales (FCEyN) de la Universidad Nacional de La Pampa (UNLPam), vinculadas con el estudio y aplicación de métodos multivariados discriminantes y de clasificación (algunos que podrían entenderse como clásicos y de una esencia más estadística, y otros propios de la minería de datos) con el propósito de establecer similitudes y diferencias, y analizar las estimaciones que se obtienen con ellos al aplicarlos efectivamente en el Análisis de Datos Multivariados (ADM).

De las investigaciones realizadas, surge el campo de la educación como un terreno propicio para las aplicaciones de Minería de Datos (MD), dada la multiplicidad de fuentes de datos y los diversos grupos de interés implicados. Asimismo, el área educativa ofrece la posibilidad de aplicar elementos de la Teoría de Respuesta al Ítem (TRI) y el Análisis de Supervivencia (AS) para el análisis de las respuestas en cuestionarios, y del tiempo requerido para la aprobación de asignaturas o la graduación, respectivamente.

En consecuencia, se inicia en 2018 un nuevo Proyecto, acreditado y financiado por la Institución mencionada, cuyo objetivo general es estudiar y aplicar distintos métodos que ofrece la MD, el ADM, la TRI y el AS, sobre los datos registrados en SIU Guarani de la FCEyN (UNLPam) con el propósito de caracterizar la trayectoria académica de los

estudiantes, y detectar patrones compatibles con situaciones de dificultades en el aprendizaje, que puedan derivar en abandono de los estudios universitarios.

1. INTRODUCCIÓN

En la actualidad, la mayoría de los procesos (industriales, académicos, negocios, servicios, entre otros) cuentan con información histórica almacenada. El avance de la tecnología ha permitido generar volúmenes de datos cada vez más grandes y difíciles de analizar y comprender. Distintas áreas han tratado de dar soluciones a este problema. La MD, en combinación con el ADM, reúnen un conjunto de técnicas capaces de modelizar y resumir la información, facilitando su comprensión y ayudando a la toma de decisiones en situaciones futuras (Cabena et al. 1998; Hernández Orallo et al., 2004). El área educativa no escapa a esta realidad. En general, los establecimientos educativos disponen de información sumamente detallada de cada alumno proveniente de múltiples fuentes (e.g. bases de datos, aplicaciones web, entornos virtuales de enseñanza y aprendizaje) pero carecen de modelos que les permitan describir de forma objetiva a sus estudiantes. Caracterizar a los estudiantes de una institución académica aporta información no trivial y de utilidad para la toma de decisiones.

La comunidad universitaria en su conjunto se plantea y propone la mejora continua de la calidad de los procesos educativos que se desarrollan en sus instituciones y de los servicios que ofrecen. La FCEyN (UNLPam) no es ajena a esta realidad. El equipo de gestión, cuerpo docente y agrupaciones estudiantiles, a través de la Comisión *ad hoc* de Ingreso y Permanencia (CIP), han diagnosticado altos niveles de deserción y desgranamiento en los primeros años de estudio. No obstante, los diagnósticos realizados carecen de la sistematización necesaria que permita revelar a tiempo el abandono de estudiantes en diferentes tramos de las carreras elegidas.

En el proceso de inscripción a las carreras de grado de la FCEyN (UNLPam), y en el

desarrollo de las actividades del Programa de Ambientación a la Vida Universitaria (PAVU) de la Institución, se recolectan múltiples datos aportados por los aspirantes a través de los sistemas de gestión que luego son enriquecidos con datos relativos a la historicidad académica de los estudiantes. Éstos constituyen una importante fuente de información, en la medida que se extraiga conocimiento para el análisis de la realidad de los estudiantes y los contextos en los que ellos aprenden, y para el diseño de eventuales planes de acción.

Este tipo de estudios en el campo de la educación corresponde a aplicaciones de una rama particular de la MD conocida como Minería de Datos Educativos (MDE). Este nuevo área de investigación interdisciplinaria se ocupa del desarrollo de métodos para explorar los datos que se dan en el ámbito educativo, así como de la utilización de estos métodos para entender mejor a los estudiantes y los contextos en que ellos aprenden (Romero & Ventura, 2010). Romero et al. (2010) definen la MDE como el desarrollo, investigación y aplicación de métodos computacionales para detectar patrones en grandes conjuntos de datos educativos que, de otro modo, serían difíciles o imposibles de analizar debido a su volumen.

Revisiones de investigaciones realizadas en MDE dan cuenta de los objetivos perseguidos y las diversas aplicaciones posibles en el área (Romero & Ventura, 2007, 2010; Baker & Yacef, 2009). Romero & Ventura (2010), en base a estas revisiones, elaboran una taxonomía de las áreas de aplicación de MDE, entre las que se menciona la predicción del desempeño de estudiantes.

Sin embargo, el estudio del rendimiento académico y del abandono escolar no es de interés reciente, y siempre ha estado relacionado con factores sociales, económicos y psicológicos. Varios estudios han abordado estos temas usando distintas metodologías: análisis discriminante, reglas de asociación, modelos de regresión logística y de imputación múltiple, ANOVA, árboles de decisión, redes neuronales, redes bayesianas, entre otros (Streeter & Franklin, 1991; Ma et

al., 2000; Wayman, 2001; Pursley, 2002; Minaei-Bidgoli et al., 2003; Kotsiantis et al., 2004; Pardos et al., 2006; Cortez & Silva, 2008; Márquez Vera et al., 2012).

Asimismo, el área educativa ofrece la posibilidad de aplicar elementos de la TRI y el AS. En particular, la TRI ofrece estimaciones del rasgo latente de individuos medidos mediante un test o cuestionario. Su utilidad en el campo educativo radica en determinar si un estudiante consigue responder correctamente a cada una de las preguntas que componen el cuestionario y en atender al puntaje bruto obtenido en la prueba (Bartholomew & Knot, 1980; Bartholomew et al., 2008; Burga León, 2005; Debera & Nalbarte, 2006; Hidalgo Flores, 2007; Pardo Adames, 2001). El AS, por su parte, aporta técnicas que permiten extraer conclusiones del tiempo requerido para la aprobación de espacios curriculares o la graduación, así como su relación con predictores sociodemográficos y de aptitud académica, entre otros (Rojas Torres & Alfaro Rojas, 2014; Gallardo Allen et al., 2016).

En suma, la MDE mediante técnicas de MD, ADM, TRI, y AS es un área de investigación relativamente reciente y de crecimiento notable. La línea de investigación que aquí se describe pretende realizar un aporte desde el área sobre la realidad y contexto de la FCEyN (UNLPam), proporcionando modelos que permitan caracterizar la trayectoria académica de los estudiantes, y detectar patrones compatibles con situaciones de dificultades en el aprendizaje y abandono. Estos modelos podrían ser de utilidad para implementar políticas de retención adecuadas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Como se mencionó anteriormente, la línea de investigación aquí presentada se desprende de un Proyecto anterior que permitió investigar técnicas de discriminación y clasificación multivariadas, con el propósito de establecer similitudes o diferencias, y analizar la eficiencia de las mismas al aplicarlas en el análisis de datos multivariados.

Habiendo identificado el campo de la educación como un terreno propicio para las aplicaciones de muchas de las técnicas estudiadas, en esta nueva línea se pretende estudiar y aplicar distintos métodos que ofrecen la MD y el ADM, la TRI y el AS, sobre los datos registrados en el sistema de gestión de información estudiantil (SIU Guarani) de la FCEyN (UNLPam) con el propósito de caracterizar la trayectoria académica de los estudiantes, y detectar patrones compatibles con situaciones de dificultades en el aprendizaje, que puedan derivar en abandono de los estudios universitarios. Los resultados obtenidos serán evaluados y comparados de manera que los mejores modelos resultantes podrían ser de utilidad en la identificación temprana de estudiantes en riesgo, y el establecimiento de una política de apoyo académico adecuada para atender la situación y, eventualmente, disminuir los índices de fracaso y abandono.

3. RESULTADOS OBTENIDOS/ESPERADOS

Debido a que la presente línea de investigación recién se inicia, no se cuenta a la fecha con resultados propios.

Hasta el momento se ha realizado una revisión sistemática de bibliografía referida a experiencias que utilicen la MDE para identificar modelos que describen la trayectoria académica de estudiantes y patrones de deserción o abandono, poniendo especial atención a las técnicas utilizadas vinculadas con el ADM, la MD, la TRI, y el AS. En particular, interesan aquellas experiencias en el ámbito de la Educación Superior de la República Argentina.

Habiendo identificado las técnicas empleadas en los estudios empíricos revisados, y considerando otras que pudieran resultar de utilidad, se espera comenzar con su aplicación sobre los datos provenientes de SIU Guarani de la FCEyN (UNLPam), previo desarrollo de técnicas de preprocesamiento (limpieza, transformación, selección de variables, y la transformación o combinación de éstas) que permitan obtener una vista minable de los datos recopilados.

Aplicadas las técnicas seleccionadas, se evaluarán y compararán los patrones y modelos resultantes a partir de un análisis e interpretación del conocimiento obtenido. Esto permitirá seleccionar los modelos más expresivos, para finalmente elaborar conclusiones pertinentes y comunicar los resultados alcanzados.

Se espera así, en un plazo no superior a los cinco años, contribuir a la identificación temprana de estudiantes en riesgo, y el establecimiento de estrategias académicas adecuadas para atender la situación y, eventualmente, disminuir los índices de fracaso y abandono.

4. FORMACIÓN DE RECURSOS HUMANOS

En la línea de investigación presentada, bajo la dirección de la Dra. Martín, y la co-dirección de la Lic. Dieser, trabajan cuatro docentes/auxiliares investigadoras (tres perteneciente a la FCEyN (UNLPam) y una de la Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur), todas con formación de base matemática. Una de ellas ha finalizado el cursado de la Maestría en Tecnología Informática Aplicada en Educación de la Facultad de Informática de la Universidad Nacional de La Plata y se encuentran en proceso de elaboración del proyecto de tesis. Las restantes han comenzado sus estudios de Doctorado en Estadística en la Universidad Nacional de Rosario, y una de ellas proyecta realizar su trabajo de Tesis Doctoral en AS, línea que, como ya se manifestara, se plantea aplicar para el estudio de la permanencia de los estudiantes universitarios.

El equipo de trabajo cuenta también con la participación de una Becaria, auxiliar docente en la FCEyN (UNLPam) quien lleva adelante su proyecto de tesis, bajo la dirección de la Dra. Martín, “La Teoría de Respuesta al Ítem aplicada a prueba diagnóstica de ingreso universitario”, a fin de obtener el grado de master en la Maestría en Estadística Aplicada de la Universidad Nacional de Córdoba. Finalmente, como asistente de investigación, se incorpora una estudiante de Licenciatura en

Matemática que ha orientado su formación específica en temas de estadística aplicada. Se espera que pueda aplicar aquí los aprendizajes apropiados, adquirir nuevos conocimientos, y eventualmente, iniciar estudios de postgrado en temas vinculados con los de este Proyecto.

5. BIBLIOGRAFÍA

Baker, R. S. J. D. & Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, 1(1):3–16.

Bartholomew, D.J. & Knott, M. (1980). *Latent Variables Models and Factor Analysis*. Kendall's Library of Statistics, 1° Edition.

Bartholomew D.J.; Steele, F., Moustaki, I. & Galbraith, J.I. (2008). *Analysis of multivariate social science data*. 2° Edition. Boca Ratón, EEUU: Taylor & Francis Group.

Burga León A. (2005). *Evaluación del rendimiento académico. Introducción a la Teoría de Respuesta al Ítem*. Ministerio de Educación, Lima. Perú.

Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., & Zanasi, A. (1998). *Discovering data mining: from concept to implementation*. New Jersey: Prentice Hall.

Cortez, P. & Silva, A. (2008). Using data mining to predict secondary school student performance. En Brito, A. and Teixeira, J. (Eds.), *Proceedings of 5th Future Business Technology Conference*, pp. 5–12, Porto, Portugal. EUROSIS.

Debera, L. & Nalbarte, L. (2006). *Pruebas diagnósticas: una aplicación a la teoría de respuesta al ítem, aproximación clásica y bayesiana*. Instituto de Estadística. F.C.E. y Administración, Universidad de la República.

Gallardo Allen, E, Molina Delgado, M. & Cordero Cantillo, R. (2016). *Aplicación del Análisis de Supervivencia al Estudio del Tiempo Requerido para Graduarse en Educación Superior: El Caso de la*

Universidad de Costa Rica. *Páginas de Educación*, 9(1):61–87.

Hernández Orallo, J., Ramírez Quintana, M. J., & Ferri Ramírez, C. (2004). *Introducción a la Minería de Datos*. Madrid: Pearson Prentice Hall.

Hidalgo Flores, R. (2007). *Teoría de respuesta al ítem: una aplicación educativa*. Facultad de Ingeniería, Universidad Autónoma de Querétaro, México.

Kotsiantis, S., Pierrakeas, C., & Pintelas, P. (2004). Predicting student's performance in distance learning using machine learning techniques. *Applied Artificial Intelligence*, 18(5):411–426.

Ma, Y., Liu, B., Wong, C. K., Yu, P. S., & Lee, S. M. (2000). Targeting the right students using data mining. En *Proceedings of 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 457–464, Boston, USA.

Márquez Vera, C., Romero Morales, C., & Ventura Soto, S. (2012). Predicción del Fracaso Escolar Mediante Técnicas de Minería de Datos. *IEEE-RITA*, 7(3):109–117.

Minaei-Bidgoli, B., Kashy, D. A., Kortemeyer, G., & Punch, W. F. (2003). Predicting student performance: an application of data mining methods with an educational web-based system. En *Proceedings of 33rd Annual Frontiers in Education, FIE 2003*, pp. 13–18, Colorado, USA.

Pardo Adames, C. (2001). El modelo de Rasch: Una alternativa para la evaluación educativa en Colombia. Facultad de Psicología. Universidad Católica de Colombia.

Pardos, Z. A., Heffernan, N. T., Anderson, B., and Heffernan, C. L. (2006). Using fine-grained skill models to fit student performance with bayesian networks. En *Proceedings of the Workshop in Educational Data Mining held at the 8th International*

Conference on Intelligent Tutoring Systems, Taiwan.

Pursley, M. (2002). *Changes in Personal Characteristics of Mexican-American High School Graduates and Dropouts During the Transition from Junior High to High School*. Texas Tech University.

Rojas Torres, L. & Alfaro Rojas, L. (2014). Análisis de sobrevivencia para la estimación del tiempo adicional como adecuación para la aplicación de una prueba estandarizada. *PEL: Pensamiento Educativo. Revista de Investigación Educativa Latinoamericana*. 51(1), 135-155.

Romero, C. & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Syst. Appl.*, 33(1):135–146.

Romero, C. & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 40(6):601–618.

Romero, C., Ventura, S., Pechenizky, M., & Baker, R. (2010). *Handbook of Educational Data Mining*. Chapman and Hall CRC Press, Taylor & Francis Group, Boca Raton.

Streeter, C. L. & Franklin, C. (1991). Psychological and family differences between middle class and low income dropouts: A discriminant analysis. *The High School Journal*, 74(4):211–219.

Wayman, J. C. (2001). Factors influencing GED and diploma attainment of high school dropouts. *Education Policy Analysis Archives*, 9(4):1–19.

Desarrollo de una Plataforma de Capacitación a Distancia Utilizable y Abierta para Personas con Discapacidad Visual

Guillermo Javier Lafuente¹, Carlos Ballesteros², José Luis Filippi³

GIAU⁴ – Facultad de Ingeniería – UNLPam.

Calle 110 esq. 9 n° 390

{lafuente¹, balleste², filippij³}@ing.unlpam.edu.ar

⁴Grupo de Investigación de Ambientes Ubicuos – <http://giau.ing.unlpam.edu.ar/>

Resumen

Este proyecto de “Inclusión en la web, como diseño Universal para personas con discapacidad visual” liderado por un equipo interdisciplinario de la Facultad de Ingeniería de la Universidad Nacional de La Pampa buscan implementar una plataforma de educación a distancia accesible y utilizable que permita brindar cursos de capacitación inclusiva a personas no videntes o con disminución visual y a todas aquellas personas que deseen utilizarla como herramienta de aprendizaje. Los mismos serán dictados por la Fundación BienEstar¹, la cual se especializa en el trabajo con personas con discapacidad.

Para contribuir a la resolución del proyecto se implementará un EAD (Entorno de Educación a Distancia) que contemple las características de Accesibilidad explicitadas en el (Artíc. 9° de la Convención Internacional de los Derechos de las Personas con Discapacidad Ley (26.378) y que sirva como plataforma de capacitación en línea. Además, se elaborará un Recurso Educativo Abierto y Accesible (REAA),

¹ Fundación BienEstar, Realicó, La Pampa
<http://www.fundacionparaelsenbienestar.org/>

como elemento capacitador en alguna disciplina que sea de interés para la fundación BienEstar.

Palabras clave: Inclusión Social, Accesibilidad, Educación a Distancia, Discapacidad Visual.

Contexto

Este proyecto se lleva a cabo en el marco de las líneas de I+D desarrolladas por el grupo GIAU (Grupo de Investigación de Ambientes Ubicuos) de la Facultad de Ingeniería de la UNLPam.

Introducción

Este proyecto trata el diseño e implementación de una plataforma de capacitación a distancia utilizable y abierta para personas con discapacidad visual. El concepto de diseño universal utilizado en varios campos y acuñado por Ronald Mace [1], [2], se refiere a la idea de diseñar productos que sean estéticos y usables para cualquier persona independientemente de su edad, habilidad y estado. Los términos más comunes utilizados para referirse a diseño universal son: simple, intuitivo, flexible, equitativo, perceptible y tolerable al error. El termino

diseño universal está muy relacionado a otros términos como accesibilidad o usabilidad [3].

Con la aparición de las nuevas tecnologías, el término accesibilidad se ha extendido a accesibilidad a las computadoras y la mayoría de los sistemas operativos incluyen innovadoras soluciones para personas con discapacidad. Además, con el crecimiento de internet, hay una sección específica dentro de la accesibilidad a computadoras que trata la accesibilidad en la Web. Algunos autores [4] [5] han escrito acerca de este tópico, describiendo tecnologías que asisten la navegación web: reconocimiento de voz, magnificación de pantalla y lectores de pantalla. En 1999 la Web Accessibility Initiative (WAI) publicó las guías para contenido web accesible WCAG [6], [7], para mejorar la accesibilidad de la web para personas con discapacidad [8].

Situación actual del problema

Con la aparición de los dispositivos móviles, la interacción humano - computadora ha cambiado significativamente, apareciendo nuevas técnicas para evaluación de usabilidad [9], [10]. Ha ocurrido un cambio radical en el desarrollo de pantallas táctiles basadas en dispositivos móviles. En poco tiempo la interacción basada en gestos se ha convertido en estándar en muchos dispositivos móviles. Las pantallas táctiles proveen una gran flexibilidad y acceso directo a controles e información, aunque por otro lado lo hacen menos accesibles para usuarios ciegos o con impedimentos visuales.

Dado que la mayoría de las aplicaciones están diseñadas para usuario con visión, las características de accesibilidad no siempre son adecuadas para obtener resultados aceptables. Una aplicación

diseñada para usuarios videntes, con una capa extra que incluya características de accesibilidad no ha mostrado ser una opción utilizable. El usuario con discapacidad visual puede utilizar esta aplicación, pero la interfaz de usuario no ha sido concebida para ciegos, lo que concluye que una buena experiencia de usuario para el caso de ciegos, no está para nada asegurada. De acuerdo a esta premisa, se necesitan aplicaciones específicas para usuarios con baja visión y ciegos para obtener la mejor experiencia de usuario posible, un desarrollo centrado en el usuario (DCU) ciego o disminuido visual.

Fundamentación

Existen diversas barreras a las cuales se enfrentan las personas con discapacidad visual, no sólo físicas, sino también de acceso a la información, quedando muchas veces excluidas del sistema socioeducativo. Tal como plantea Piñeros [11], la desventaja radica, entonces, no en la discapacidad en sí, sino en el ámbito social, que no logra integrar a la persona discapacitada.

Los usuarios con discapacidad visual *“tienen las mismas necesidades de información que el resto de los ciudadanos”*. Estas personas deben recibir información accesible, que les permita tomar decisiones y realizar una vida independiente” [11].

La Facultad de Ingeniería de la UNLPam, busca dar respuesta a la necesidad que en la actualidad presenta la “Fundación para el BienEstar” brindando tecnologías como ámbitos de acceso, difusión de la información y servicios a la comunidad. Se intenta sentar las bases de inclusión digital en una sociedad, que desafíe las diferencias, que profundice los vínculos y que permita alcanzar mayor igualdad social y educativa para personas con

discapacidad a través de un diseño para todos con las configuraciones de apoyo tecnológico que deban adoptarse.

Atento a esto, es que se diseña una plataforma de educación a distancia para capacitación a personas con disminución visual, no videntes y demás beneficiarios que quieran hacer uso de la misma, empleando para ello, diferentes mediadores didácticos que orienten a la utilización del mismo.

En la primera conferencia de E-learning hacia la inclusión social se declararon los siguientes términos: *“El e-learning no ha de limitarse a ser cursos en línea para universidades y grandes compañías. No ha de estar centrado exclusivamente en cómo aumentar los beneficios. Cuando se desarrollen módulos de e-learning, se deberá estar seguro de que todos los grupos sociales tienen acceso a las técnicas, y darle a todo el mundo los medios para usar las TIC en su desarrollo profesional y personal, y así poder aprender en la sociedad de la información”* [12]. A través de esta línea de trabajo, los investigadores entienden que deben provocar un impacto social positivo, que tenga mayor conciencia social inclusiva; utilizando buenas prácticas que respeten la diversidad en un marco de igualdad.

Metodología de Trabajo

Desde el punto de vista del desarrollo, el proyecto será afrontado bajo el método OpenUP/Basic [13], un proceso unificado que incorpora técnicas ágiles ya probadas en la industria de software y constituye un proceso estructurado, robusto, eficiente y liviano que se adapta a los requerimientos de este proyecto.

Desde el punto de vista de la accesibilidad sobre los artefactos a implementar, se utilizarán dos estrategias para realizar la validación y verificación

de la accesibilidad teniendo en cuenta principalmente dos tipos de usuario, el no vidente y el disminuido visual. Se espera implementar dos estrategias de evaluación automatizada para verificar la accesibilidad como así también una estrategia de evaluación centrada en el usuario.

a) Estrategia de evaluación automatizada de la accesibilidad

Cada artefacto de software que se desarrolle será evaluado para la accesibilidad usando un conjunto de herramientas automatizadas [14], [15], [16]. La aplicación de las mismas, generará una completa información la cual brindará ayuda para lograr un mayor grado de accesibilidad sobre los productos analizados. Estas herramientas examinarán la accesibilidad bajo las directrices de accesibilidad web, como las WCAG, e informarán sobre las barreras conocidas y probables o potenciales (en forma de un informe independiente y / o anotaciones del producto evaluado).

Para el caso de los Recursos Educativos Abiertos Accesibles (REAAs), y que dispongan de componentes como presentaciones en powerpoint, se utilizará la característica de evaluación de accesibilidad incorporada de PowerPoint para examinar las diapositivas de PowerPoint de cada REAA en busca de posibles barreras de accesibilidad. También se trabajará con lectores de pantalla para detectar problemas de interpretación de texto en la lectura de documentos de distinta fuente o formato. Los datos obtenidos de las herramientas [14], [15], [16] serán agrupados y tabulados para preparar una lista única de problemas de accesibilidad potenciales para cada sección de los artefactos de software. Los datos obtenidos de la herramienta de evaluación de accesibilidad de PowerPoint y los

resultados de los lectores de pantalla, serán agrupados y tabulados para preparar una lista única de posibles problemas de accesibilidad para cada unidad de información contenida en los REAAs.

b) *Estrategia de evaluación de accesibilidad centrada en el estudiante/usuario*

Se seleccionará un grupo definido de estudiantes. El conjunto de estudiantes deberá estar integrado, la mitad por estudiantes sin problemas y la otra mitad con estudiantes no videntes y disminuidos visuales. Los expertos de la Fundación para el Bienestar colaborarán en esta tarea de selección. Se dispondrá de laboratorio de pruebas y software para que los participantes completen algunas actividades de un REAA. El laboratorio consistirá de una estación de trabajo en un aula universitaria de acuerdo con un formato tradicional de laboratorio de pruebas de usabilidad como los propuestos por Nielsen [17] y Rubin y Chisnell [18]. La estación de trabajo estará constituida por un escritorio con una PC, Internet por cable, Mouse y micrófono externo. Una cámara de video en un trípode estará disponible delante de la estación para capturar las expresiones faciales del estudiante y las interacciones con la PC, y el aula será visible desde una ventana vidriada para realizar una observación adyacente. Cada estudiante deberá completar la actividad en una sesión moderada (en presencia de un investigador y una cámara de video). Luego realizará una actividad en una sesión no modificada (trabajando solo en el aula con la cámara de video apagada). Durante las sesiones moderadas, el investigador deberá indicar al estudiante que piense en voz alta si es necesario siguiendo una forma de comunicación del habla del protocolo de pensar en voz alta y lo registrará [19].

Para relevar estas pruebas de laboratorio se utilizarán herramientas para captura de escenas como OpenVULab o Screencast-O-Matic², las cuales permiten crear screencasts de las actividades en pantalla realizada por los estudiantes mientras completan las consignas propuestas. Los screencasts serán sincronizados con las verbalizaciones de pensamiento en voz alta realizada por los participantes con el fin de registrar todos los detalles de la interacción con los softwares y contenidos desarrollados. Toda la actividad de laboratorio será documentada por los investigadores participantes del proyecto junto con personal idóneo propuesto por la Fundación para el Bienestar.

Líneas de Investigación, Desarrollo e Innovación

A continuación, se detallan sintéticamente las líneas de investigación en las cuales se avanzará para llevar a cabo el objetivo del proyecto.

- Entornos de Educación a Distancia, REAA.
- Usabilidad, Diseño para usabilidad, DCU, UX.
- Evaluación de usabilidad para ciegos con entornos existentes.
- Métodos, técnicas y herramientas disponibles para cada etapa del proceso DCU.

Resultados y Objetivos

Desarrollar una plataforma de educación a distancia abierta, accesible y utilizable, y un REAA a través de un diseño centrado en usuarios con discapacidad visual, que esté diseñada y

² OpenVULab (<http://openvulab.org>) Screencast-O-Matic (<https://screencast-o-matic.com/>)

probada por usuarios con discapacidad visual, involucrados en el proceso de desarrollo.

Formación de Recursos Humanos

Actualmente, el proyecto cuenta con la siguiente conformación de grupo de trabajo: un Director de Proyecto, un co-director y tres Investigadores, un estudiante becario Tesista y 4 estudiantes (3 de carreras relacionadas a informática en el ámbito de la Facultad de Ingeniería de la UNLPam., y una estudiante de la carrera Licenciatura en Ciencias de la Educación de la Facultad de Ciencias Humanas de la UNLPam). La conformación del grupo docentes afectado al proyecto ha sido seleccionado siguiendo un criterio interdisciplinar, teniendo en cuenta a docentes especialistas en Educación Virtual, Tecnología y Educación Especial entre otras, los cuales pertenecen a la Facultad de Ingeniería y a la Facultad de Ciencias Humanas de la UNLPam. Se ha incorporado una estudiante de la Facultad de Ciencias Humanas de la UNLPam, quien es disminuida visual, y será uno de los actores de mayor relevancia para alcanzar los objetivos del proyecto. Un becario tesista de la Carrera Ingeniería en Sistemas, participa con el desarrollo de tesis denominado: “*Investigación y Desarrollo de Framework para la generación de contenido Web Accesible*”.

Referencias

- [1] Ronald Mace, “The RL Mace Universal Design Institute.” [Online]. Available: <http://www.udinstitute.org/>.
- [2] R. L. Mace, “What is Universal Design (DO NOT CITE!),” 2005.
- [3] M. F. Story, “Maximizing Usability: The Principles of Universal Design,” *Assist. Technol.*, 1998.
- [4] H. Takagi, C. Asakawa, K. Fukuda, and J. Maeda, “Accessibility designer: visualizing usability for the blind,” in *ACM SIGACCESS Accessibility and Computing*, 2004, no. 77–78, pp. 177–184.
- [5] J. P. Bigham, A. C. Cavender, J. T. Brudvik, J. O. Wobbrock, and R. E. Lander, “WebinSitu,” in *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility - Assets '07*, 2007, p. 51.
- [6] S. Leuthold, J. A. Bargas-Avila, and K. Opwisa, “Beyond web content accessibility guidelines: Design of enhanced text user interfaces for blind internet users,” *Int. J. Hum. Comput. Stud.*, 2008.
- [7] “Web Accessibility Initiative (WAI) - home page | Web Accessibility Initiative (WAI) | W3C.” [Online]. Available: <https://www.w3.org/WAI/>.
- [8] J. Mankoff, H. Fait, and T. Tran, “Is your web page accessible?,” in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '05*, 2005.
- [9] J. Kjeldskov and J. Stage, “New techniques for usability evaluation of mobile systems,” *Int. J. Human-Computer Stud.*, 2004.
- [10] A. Kaikkonen, A. Kekalainen, M. Cankar, T. Kallio, and A. Kankainen, “Usability testing of mobile applications: A comparison between laboratory and field testing,” *J. Usability Stud.*, 2005.
- [11] I. Piñeros, *El acceso a la información de las personas con discapacidad visual: modelo de servicio para bibliotecas públicas*. Alfagrama Ediciones, 2008.
- [12] G. Apostopoulou, L. Baronio, and et all., “E-learning hacia la inclusión social,” pp. 1–8, 2004.
- [13] Eclipse Foundation, “Eclipse Process Framework (EPF),” *OpenUP*, 2012. .
- [14] “WAVE Web Accessibility Tool.” [Online]. Available: <http://wave.webaim.org/>.
- [15] “IDI Web Accessibility Checker: Web Accessibility Checker.” [Online]. Available: <https://achecker.ca/checker/index.php>.
- [16] “Home: AInspector Sidebar.” [Online]. Available: <http://ainspector.github.io/>.
- [17] J. Nielsen, *Usability Engineering*, vol. 44, no. 3. 1993.
- [18] D. Rubin, J., & Chisnell, *Handbook of usability testing [electronic resource]: How to plan, design, and conduct effective tests (2nd ed.)*. 2008.
- [19] M. T. Boren and J. Ramey, “Thinking aloud: Reconciling theory and practice,” *IEEE Trans. Prof. Commun.*, vol. 43, no. 3, pp. 261–278, 2000.

MODELO DE SISTEMA TUTOR INTELIGENTE DE APOYO EN LA REALIZACIÓN DE PROYECTOS DE INVESTIGACIÓN EN POSGRADOS EN INGENIERÍA.

Jaquelina E. Escalante, Sonia I. Mariño, Susana Marchisio, M. Viviana Godoy G.

Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura.
9 de Julio 1449. CP: 3400. Corrientes. Argentina.

Facultad de Humanidades. Las Heras 727. Resistencia. Chaco
Universidad Nacional del Nordeste.

Facultad de Ciencias Exactas, Ingeniería y Agrimensura. Av. C. Pellegrini 250. Rosario Universidad
Nacional de Rosario.

jaquelina_escalante@hotmail.com; smarch@fceia.unr.edu.ar; simarinio@yahoo.com;
mvgodoy@exa.unne.edu.ar

RESUMEN

En este trabajo se introducen conceptos de los campos en los que incursiona esta investigación, como son: la problemática en la elaboración de Proyectos de tesis en posgrados de Ingeniería y los Sistemas Tutores Inteligentes (STI). Se describen los aspectos esenciales que se deben considerar en el modelado de un sistema tutor inteligente, más como una herramienta complementaria de la enseñanza y del aprendizaje para apoyo de los docentes tutores, modelando el conocimiento de Expertos y Tutores, con la finalidad de aumentar la calidad del aprendizaje.

Palabras clave: sistemas tutores inteligentes, metodología de la investigación, Modelo del tutor, redacción de proyectos de tesis. Ingeniería.

CONTEXTO

El proyecto “Modelado y Diseño de un Sistema Tutorial Inteligente, orientado a la Enseñanza de la Formulación de Proyectos de Investigación en Posgrados de Ingeniería” [1] se desarrolla para la finalización del Doctorado en Ingeniería de la FCEIA – UNR y el marco de la beca de Investigación de la SGCyT – UNNE y como integrante de los Proyecto de I+D: “Sistemas de Información y TIC: métodos y herramientas” [12] y “TI en los Sistemas de Información y TIC: modelos, métodos y herramientas” [13] acreditados por

la Secretaria General de Ciencia y Técnica (Universidad Nacional del Nordeste).

1. INTRODUCCIÓN

En la literatura se localizan una variedad de estudios que alertan sobre el escaso desarrollo e interiorización de actitudes y habilidades para la investigación científica en la formación universitaria [2] [3] [4] [5].

Según Carlino [4] la formación en investigación es un proceso largo y continuo; investigar requiere de formación en la práctica. Elaborar una tesis de posgrado es usualmente el camino para iniciarse en la cultura de la investigación, cuyas prácticas y valores centrales suelen aguardar a los tesisistas de forma implícita

En Marchisio [7] se menciona que los resultados que emergen del análisis de la experiencia de dictado y de apoyo tutorial en asignaturas como Metodología de la Investigación en doctorados y maestrías, y de Talleres de Tesis en carreras de maestría en la Facultad de Ciencias Exactas, Ingeniería y Agrimensura de la Universidad Nacional de Rosario (FCEIA – UNR) revelan que, al momento de elaborar su proyecto, los estudiantes manifiestan diversas dificultades.

La mediación tecnológica en esta instancia de formación facilita, además, del intercambio de experiencias, el sostenimiento de las cercanías del tesisista con la institución y los directores, y el debate interdisciplinario sobre proyectos entre pares.

En este contexto, el **DOCENTE / TUTOR**, es responsable de proporcionar el “contenido de aprendizaje”, actuando en calidad de experto en el dominio del campo de conocimiento metodológico, aportando conceptos, pero además acompaña y apoya promoviendo la emergencia y desarrollo de los procesos cognitivos complejos requeridos para alcanzar la resolución del problema con diversas estrategias.

Por otro lado, según diversas definiciones, un Sistema Tutor Inteligente (STI) “*es un sistema de software que utiliza técnicas de IA para representar el conocimiento e interactúa con los estudiantes para enseñárselo*” [8], [9], [11], [10].

Estos STI han evolucionado desde sus inicios y son un ejemplo de sistemas educativos adaptables. La adaptabilidad es una importante característica de estos sistemas, pues en lugar de presentar información estática, el conocimiento es presentado de manera personalizada y dinámica al estudiante de acuerdo a su propio comportamiento. De esta forma, cada estudiante recibe retroalimentación de acuerdo a su interacción con el sistema. Estos sistemas pueden trabajar con reglas que permiten encontrar habilidades dominadas por los estudiantes y reforzar aquellas en donde se detectan carencias en su comprensión.

Básicamente un STI consta de los siguientes componentes: ***un módulo del dominio del conocimiento***: la base de conocimiento del tutor, esto es, todos los conceptos y relaciones necesarias en el proceso de enseñanza-aprendizaje, ***un módulo del estudiante*** (que es capaz de definir el conocimiento del estudiante en cada punto durante la sesión de trabajo, ***un módulo del tutor***: que genera las interacciones de aprendizaje basadas en las discrepancias entre el especialista y el estudiante y finalmente la interface con el usuario: que permite la interacción del estudiante con un STI de una manera eficiente (conocimiento sobre cómo presentar los contenidos).

El proyecto de tesis para el “modelado y diseño de un sistema tutor inteligente, orientado a la enseñanza de la formulación de proyectos de investigación en posgrados de Ingeniería”, tiene como propósito elaborar un modelo conceptual y diseñar un Sistema Tutorial Inteligente adaptativo como arquitectura de apoyo a docentes y estudiantes de posgrado que cursando las asignaturas del área de fundamentos metodológicos de la investigación en posgrados de Ingeniería se enfrentan a la tarea o problema de formular un proyecto de investigación relevante. Esto a modo de brindar al estudiante material pedagógico sobre el tema a enseñar, para una mejor comprensión del saber investigar en el ámbito de la ingeniería, a partir del desarrollo de ese saber escribir, monitorizándolo, teniendo en cuenta el ***modelo del estudiante*** desde sus características personales que condicionarán el aprendizaje, así como las de su comportamiento y evolución. Además, deberá utilizar las estrategias pedagógicas más adecuadas para el logro de los objetivos propuestos.

Para ello deberá reconocerse, la habilidad cognitiva y preferencias de aprendizaje del estudiante según sus estilos de aprendizaje, y se deberá personalizar el ambiente de aprendizaje y su correspondiente seguimiento, con lo que se favorecerá el desarrollo de la autonomía en el proceso de aprendizaje de la formulación de proyectos de investigación en ingeniería.

En este trabajo, el diseño e implementación del STI modela el conocimiento de los profesores. Es decir, se centra en el ***DOCENTE TUTOR***, en el análisis de cómo se ha adquirido y diseñado las capacidades y conocimiento pedagógico, las ideas que debe aportar cada profesor para capacitar al estudiante en su proceso de aprendizaje o de formación de conocimientos al momento de realizar proyectos de investigación, y como asegurar que sean asimilados de una manera eficaz, eficiente, fiable y adecuada

Por lo expuesto, el STI que surge de esta propuesta, conjuntamente con los cursos de Metodología de Investigación y Talleres de

Tesis en la modalidad a distancia mediados por un entorno virtual y proporcionados por la Facultad deberán facilitar y apoyar al estudiante, es decir al tesista, para que en compañía de su director y apoyado por su tutor formule su proyecto.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Los aspectos importantes en los que se enfoca este proyecto de tesis son:

- Identificación, análisis y selección de modelos pedagógicos, considerando competencias y tipos de competencias tanto de docentes tutores como de alumnos de posgrado
- Análisis de STI disponibles en la web o descritos en las publicaciones.
- Identificación y selección de Tecnología Informática aplicada en Educación.
- Identificación y selección de Tecnología de la Inteligencia Artificial aplicada en Educación.
- Identificación, análisis y selección de modelos pedagógicos adaptables al entorno virtual.
- Diseño y modelo de un STI para integrar los aspectos teóricos y metodológicos contemplados en el proyecto.

3. RESULTADOS Y OBJETIVOS

Los STI permiten la emulación de un tutor humano para determinar qué enseñar, cómo enseñar y a quién enseñar a través de un módulo del dominio. Como resultados preliminares del proyecto se mencionan:

- Se relevan, clasifican y analizan diversos STI identificados en la web o en publicaciones.
- Se determinan y analizan modelos pedagógicos implementables en entornos virtuales, algunos de ellos se orientan al tutor y otros a los alumnos
- Se identifican y estudian los métodos de la Inteligencia Artificial que intentan simular a los especialistas
- Se diseñan y aplican encuestas para clasificar la población de alumnos de posgrados en Ingeniería de acuerdo a un perfil en relación con la

información recabada (conocimientos, habilidades, capacidades y actitudes) mediante encuestas a los cursantes de carreras de Maestría y Doctorado

- Se diseñan y realizan entrevistas a docentes y tutores a modo de determinar aquellas competencias conceptuales, metodológicas y humanas, procedimientos y actitudes a considerar para tutorizar a un alumno de tesis de posgrado.
- Se indaga en los principios del diseño, implementación y evaluación de sistemas computacionales interactivos para crear interfaces fáciles de comprender por los destinatarios y así asegurar su utilización.

4. FORMACIÓN DE RECURSOS HUMANOS

En el marco de este proyecto se desarrolla un Doctorado en Ingeniería, en la FCEIA – UNR. El conocimiento adquirido y consolidado a través de la propuesta se podrá transferir a los ámbitos académicos como soporte a los procesos de seguimiento. Además, se espera formar recursos humanos en los temas tratados.

5. BIBLIOGRAFIA

- [1] J. E. Escalante, S Marchisio, S. Mariño, “Modelado y Diseño de un Sistema Tutorial Inteligente, basado en conocimiento procedimental y cualitativo, orientado a la Enseñanza de la Formulación de Proyectos d Investigación en Posgrados de Ingeniería” Plan de tesis aprobado Doctorado en Ingeniería. Facultad de Ciencias Exactas, Ingeniería y Agrimensura. UNR, 2015.
- [2] C. Ambrosini, A. Mombrú Ruggiero, “Bendita tesis: tribulaciones personales e institucionales” *Perspectivas Metodológicas* N°19. Vol. II, 2017.
- [3] J.I. Pérez Rave. “El artículo de revisión sistemática como vehículo de escritura, investigación y publicación en ingeniería”. *Investigación y Ciencia*

- de la Universidad Autónoma de Aguascalientes. Número 64: 70-77, enero-abril 2015.
- [4] P. Carlino. “Alfabetización académica: Un cambio necesario, algunas alternativas posibles”. EDUCERE: La Revista Venezolana de Educación, 6(20): 409-420, 2003.
- [5] M. V. Díaz Alarcón “La formación para la investigación de los docentes universitarios como agentes de cambio ante los nuevos desafíos”. Didasc@lia: Didáctica y Educación, III(1): 13- 24, 2012.
- [6] N. Sotero Rodríguez León. Tesis doctoral: “Modelo de Formación del tutor de Tesis” Facultad de Educación de la Universidad Nacional de educación a Distancia – UNED, 2010.
- [7] P. Del Rio, S. Ferrara, S. Marchisio, S. Concari. “Metodología de la Investigación a distancia. Mejora continua en posgrados de Ingeniería”. Congreso Latinoamericano de Ingeniería (CLADI 2017). 13, 14 y 15 de septiembre de 2017.
- [8] J. R. Anderson, M. Matessa, C. Lebiere, “ACT-R: A theory of higher level cognition and its relation to visual attention”. *Human-Computer Interaction*, 12(4), 439-462, 1997.
- [9] L. M. M. Giraffa, M. A. Nunes, and R.M. Viccari, “Multi-Ecological: a Learning Environment using Multi-Agent architecture”. *Proc. of MASTA'97 (EPIA'97 workshop on Multi-agent Systems: Theory and Applications)*, Coimbra, Brasil, 1997.
- [10] A. Gordillo Guillen, H. Andrade Gómez, R. Rivera López. “Modelo de un sistema tutor inteligente para el desarrollo del pensamiento computacional” *Revista Investigación Educativa. Journal CIM* Vol. 5, Núm. 2, pp. 1479 – 1485. 2017.
- [11] K. VanLehn, *Student Modelling*, en M.C. Polson, and J.J. Richardson, (Eds.) *Foundations of Intelligent Tutoring systems*. Taylor & Francis, 2013.
- [12] S. I. Mariño, María V. Godoy G. “Sistemas de Información y TIC: métodos y herramientas” Proyecto de I+D, acreditado por la Secretaria General de Ciencia y Técnica, Universidad Nacional del Nordeste, 2013.
- [13] S. I. Mariño, M. V. Godoy G. “TI en los Sistemas de Información y TIC: modelos, métodos y herramientas”, Proyecto de I+D, acreditado por la Secretaria General de Ciencia y Técnica, Universidad Nacional del Nordeste, 2016.

Uzi: Máquina virtual sobre plataforma Arduino para robótica educativa

Resultados de la primer etapa

Ricardo Moran, Gonzalo Zabala, Matías Teragni, Sebastián Blanco

Centro de Altos Estudios en Tecnología Informática
Facultad de Tecnología Informática
Universidad Abierta Interamericana
Av. Montes de Oca 745, Ciudad Autónoma de Buenos Aires, República Argentina
(+54 11) 4301-5323; 4301-5240; 4301-5248
{Ricardo.Moran, Gonzalo.Zabala, Matias.Teragni, Sebastian.Blanco}@uai.edu.ar

Resumen

El objetivo de este proyecto es el desarrollo de una máquina virtual para Arduino que ejecute un set de instrucciones definido especialmente para facilitar la enseñanza de robótica en las escuelas.

Palabras clave: Physical Etoys, Arduino, máquina virtual, lenguaje de programación, robótica educativa

Contexto

El presente proyecto será radicado en el Centro de Altos Estudios en Tecnología Informática (CAETI), dependiente de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana. El mismo se encuentra inserto en la línea de investigación “Sociedad del conocimiento y Tecnologías aplicadas a la Educación”. El financiamiento está dado por la misma Universidad Abierta Interamericana

Introducción

En los últimos años, la aparición de kits de robótica orientados a usuarios no expertos fomentó el desarrollo de un conjunto significativo de proyectos educativos usando robots en diferentes niveles de educación, desde jardín de infantes hasta educación de grado. La plataforma de hardware Arduino, siendo abierta y de bajo costo, se ha popularizado muy rápidamente a nivel mundial para proyectos de este estilo.

Arduino provee un entorno de desarrollo simplificado (basado en el lenguaje de programación C++) en el que muchos conceptos avanzados de programación de microcontroladores están “escondidos” al usuario. Sin embargo, este entorno es todavía demasiado complejo para algunos de los usuarios menos experimentados, sobre todo los niños más pequeños. Esto limita la complejidad de los problemas que estos usuarios pueden atacar.

Por estas razones, y aprovechando el carácter abierto de Arduino, han surgido

múltiples intentos de proveer un entorno de programación más adecuado para principiantes. Uno de estos intentos es Physical Etoys, un ambiente de programación visual para chicos diseñado para proveer mecanismos de comunicación con distintas plataformas de hardware, incluyendo Arduino. Usando Physical Etoys, un alumno puede programar robots de manera completamente gráfica, simplemente arrastrando y soltando instrucciones en la pantalla.

Sin embargo, Physical Etoys tiene un problema: dado que los programas se ejecutan en la computadora y las órdenes son transmitidas hacia el robot, Physical Etoys requiere que los robots estén continuamente conectados a la computadora para poder funcionar. Si bien la interactividad que este modelo de programación provee es muy importante (sobre todo para los alumnos que recién están comenzando), la falta de autonomía es un limitante muy importante. Por un lado, la comunicación constante con la computadora no está permitida en algunas competencias. Por otro lado, la cantidad de proyectos que pueden realizarse de esta forma es limitada (por ejemplo, quedan fuera algunas actividades de robótica situada que requieran la toma de decisiones en un entorno complejo y dinámico). Otros entornos de programación similares a Physical Etoys presentan el mismo problema (por ejemplo, Scratch 4 Arduino).

Una posible solución sería compilar los programas generados visualmente a un código nativo que pueda ejecutar el robot sin necesidad de una conexión con la computadora. Esta opción es la que eligieron otros entornos de programación

(por ejemplo, Minibloq). Sin embargo, esta solución tiene sus propios problemas. Separar el “tiempo de compilación” del “tiempo de ejecución” imposibilita la forma de trabajo interactiva que caracteriza a ambientes como Physical Etoys.

Líneas de Investigación, Desarrollo e Innovación

No existe un entorno de programación visual para robótica que resuelva estos problemas, por lo tanto, creemos necesario un enfoque diferente. Los siguientes requerimientos deben cumplirse:

1. La ejecución de los programas debe hacerse directamente en el Arduino sin necesidad de interacción con la computadora (a excepción de la transmisión del programa).
2. En caso que el Arduino estuviera conectado con la computadora, todas las posibilidades de interacción que ofrece Physical Etoys deben mantenerse.
3. Para el usuario final debe ser transparente si el arduino funciona en modo “autónomo” o “interactivo”.

Otros requerimientos importantes (aunque no esenciales):

4. El entorno de programación debe ofrecer mecanismos de abstracción que permitan ocultar los detalles de algunos conceptos avanzados de programación de microcontroladores (por ejemplo: timers, interrupciones,

modelos de concurrencia, etc.) Estos conceptos pueden luego ser introducidos a un ritmo compatible con las necesidades de los alumnos, pero desconocerlos no debería ser un limitante.

5. El entorno de ejecución debe ofrecer mecanismos para encontrar y arreglar errores (debugging). Particularmente, debe ser posible detener en cualquier momento la ejecución, observar el estado interno del programa, y continuar ejecutando paso a paso de forma interactiva.

Nuestra propuesta para cumplir con estos requisitos es la utilización de una máquina virtual que ejecute en el Arduino un set de instrucciones sencillo diseñado especialmente para robótica educativa. Esta máquina virtual será transmitida (compilada) al Arduino una sola vez, y a partir de ese momento las herramientas de desarrollo diseñadas especialmente para este propósito podrán comunicarse con la máquina virtual para transmitir instrucciones individuales o programas enteros a través del puerto serie.

Esta máquina virtual, además de ejecutar los programas, será la encargada de abstraer todo mecanismo de comunicación y de detectar cuando el Arduino está conectado con la computadora, en cuyo caso podrá comunicarse con la misma en forma interactiva (ya sea enviando información de los sensores y estado interno del motor de ejecución, o recibiendo comandos que modifiquen el estado de los actuadores).

Si bien no se conocen entornos que cumplan con los requisitos antes planteados, el desarrollo de máquinas

virtuales y lenguajes de programación de alto nivel para microcontroladores como Arduino no es nuevo. La mayoría de los desarrollos relevados se basan en lenguajes de programación de propósito general como Java, Scheme, o Python.

Resultados y Objetivos

En esta primera etapa hemos avanzado en el diseño de un conjunto de instrucciones para robótica educativa así como el desarrollo de un prototipo de la máquina virtual encargada de ejecutar estas instrucciones. Tomando como base este desarrollo hemos definido un lenguaje de programación de alto nivel con una sintaxis inspirada en C para el cual hemos desarrollado un compilador que permite generar las instrucciones que puede entender la máquina virtual. Para facilitar el desarrollo hemos desarrollado además un editor de código simplificado.

Los objetivos para la siguiente etapa serán:

1. Validar el diseño del lenguaje y del conjunto de instrucciones usando la máquina virtual prototípica.
2. Optimizar el uso de recursos tanto de la máquina virtual como del compilador.
3. Avanzar en el desarrollo del editor para incluir herramientas de depuración de código.
4. Integrar el desarrollo con Physical Etoys

Formación de Recursos Humanos

El equipo de trabajo está conformado por un investigador adjunto del Centro de Altos Estudios en Tecnología Informática (CAETI) quien ejerce el rol de director del proyecto, dos doctorandos, y un

ayudante alumno de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana.

Bibliografía

- J.E. Smith, R. Nair: “Virtual Machines: Versatile Platforms for Systems and Processes” in The Morgan Kaufmann Series in Computer Architecture and Design Series, Morgan Kaufmann Publishers (2005)
- C. Schlegel, “Communication patterns as key towards component interoperability,” in Software Engineering for Experimental Robotics (Series STAR, vol. 30), D. Brugali, Ed. Berlin, Heidelberg: Springer-Verlag, , pp. 183–210. Smartsoftware (2007)
- Ando, N., Suehiro, T., Kitagaki, K., Kotoku, T., Yoon, W.K.: RT-middleware: Distributed component middleware for RT (robot technology). In: International Conference on Intelligent Robots and Systems 2005 (IROS 2005), pp. 3933–3938 (2005)
- Brooks, A., Kaupp, T., Makarenko, A., Oreback, A., Williams, S.: Towards component-based robotics. In: Proc. of 2005 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS’05), pp. 163–168. Alberta, Canada (2005)
- Christian Schlegel, Thomas Haßler, Alex Lotz and Andreas Steck. Robotic Software Systems: From Code-Driven to Model-Driven Designs. In procs. Of ICAR 2009. International Conference on Advanced Robotics. IEEE Press (2009)
- Bert Freudenberg, Yoshiki Ohshima, and Scott Wallace, "Etoys for one laptop per child," in 7th International Conference on Creating, Connecting and Collaborating through Computing - C5 2009, Kyoto, 2009, pp. 57-64.
- R. Rahul, A. Whitchurch, and M. Rao, "An open source graphical robot programming environment in introductory programming curriculum for undergraduates," in 2014 IEEE International Conference on MOOCs, Innovation and Technology in Education, IEEE MITE 2014, Patiala, 2014, pp. 96-100.
- A. Rao, "The application of LeJOS, Lego Mindstorms Robotics, in an LMS environment to teach children Java programming and technology at an early age," in 5th IEEE Integrated STEM Education Conference, ISEC 2015, 2015, pp. 121-122.
- M. Elizabeth and C. Hull, "Occam-A programming language for multiprocessor systems," Computer Languages, vol. 12, no. 1, pp. 27-37, 1987.
- C. L. Jacobsen, M. C. Jadud, O. Kilic, and A. T. Sampson, "Concurrent event-driven programming in occam- π for the Arduino," Concurrent Systems Engineering Series, vol. 68, pp. 177-193, 2011.
- Ryan Suchocki and Sara Kalvala, "Microscheme: Functional programming for the Arduino," in Scheme and Functional Programming Workshop, Washington, D.C., 2014, pp. 21-29.

ANÁLISIS DE ASPECTOS DE DISEÑO EN CURSOS MASIVOS EN LÍNEA CASO DE ESTUDIO SOBRE UN CURSO DE ACCESIBILIDAD WEB

Javier Díaz, Alejandra Schiavoni, Paola Amadeo, Ivana Harari,
Alejandra Osorio, Federico Carrilao Ávila
LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.

Facultad de Informática. Universidad Nacional de La Plata
Calle 50 esq. 120, 2do Piso. Tel: +54 221 4223528

jdiaz@unlp.edu.ar, ales@info.unlp.edu.ar, pamadeo@linti.unlp.edu.ar, iharari@info.unlp.edu.ar,
aosorio@cespi.unlp.edu.ar, federicoca95@gmail.com

RESUMEN

El presente artículo presenta las líneas de investigación que se vienen llevando a cabo en el Laboratorio de Investigación en Nuevas Tecnologías Informática de la Facultad de Informática de la UNLP, en relación a los ambientes de aprendizaje y al concepto de aprendizaje abierto. El aprendizaje abierto es una concepción de educación que tiene por objetivo eliminar todas las barreras del aprendizaje y proporcionar al estudiante un sistema de educación y entrenamiento centrado en sus necesidades específicas y localizado en múltiples áreas de conocimiento. Dentro de este ámbito se consideran conceptos vinculados a recursos educativos abiertos, OER (Open Educational Resources) y materiales de cursos abiertos, OCW (OpenCourseware). En relación a esto es que surgieron los MOOCs como consecuencia de las nuevas tendencias internacionales en el aprendizaje basado en la tecnología, como parte del movimiento educativo abierto.

Continuando con lo planteado en WICC 2017, en esta etapa las líneas a seguir parten de la experiencia del caso de uso desarrollado sobre Accesibilidad Web. Comprenden comparativas de metodologías de diseño, evaluación de la plataforma utilizada y estudio de resultados de la experiencia llevada a

cabo en la implementación de un curso MOOC de accesibilidad Web, con el objetivo de facilitar el proceso de creación de este tipo de cursos.

Palabras clave: Recursos educativos abiertos, OER, MOOC, accesibilidad

CONTEXTO

El proyecto descripto en este artículo se desarrolla en el Laboratorio de Investigación en Nuevas Tecnologías Informáticas, LINTI de la Facultad de Informática de la UNLP y está enmarcado en el proyecto 11-F020 “Internet del futuro: Ciudades digitales inclusivas, innovadoras y sustentables, IoT, ciberseguridad y espacios de aprendizaje del futuro”, acreditado en el marco del Programa de Incentivos, bajo la dirección del Lic. Javier Díaz.

La Facultad de Informática tiene una amplia experiencia en el uso de la plataforma Moodle para la creación de cursos que se utilizan como complemento de las clases presenciales y en cursos completamente no presenciales y semipresenciales desde el año 2003. A partir de esto, y considerando la tendencia actual de brindar cursos masivos, es que se comenzó a desarrollar un MOOC sobre Accesibilidad Web.

El tema de accesibilidad se viene trabajando en la facultad desde el año 2002, a lo largo de los cuales se incorporó la temática en el plan de estudios de la asignatura Diseño Centrado en el Usuario, se institucionalizó su abordaje mediante la creación de una Dirección de Accesibilidad, se desarrollaron tesinas, trabajos de cátedra al respecto y diferentes proyectos de extensión acreditados por la Universidad Nacional de La Plata. Los proyectos de los últimos dos años son “Trabajando por una Web Accesible” [1] y “Trabajando por una Web Inclusiva: un desafío que nos compromete a todos” [2].

1. INTRODUCCIÓN

Los OER ayudan a proporcionar apoyo al aprendizaje abierto, que es en sí una concepción de educación que tiene por objetivo eliminar todas las barreras del aprendizaje y permitir que el proceso de aprendizaje se centre en el estudiante, y en sus necesidades específicas.

Actualmente se está dando un proceso de cambio en la educación basado en los nuevos modelos de e-Learning y el desarrollo de la tecnología. En esta etapa el surgimiento de los MOOCs representa una manifestación muy evidente y de gran recorrido, convirtiéndose en protagonistas importantes en la educación en el año 2012 [3], aunque existieron algunas experiencias anteriores. Los MOOCs surgieron como consecuencia de las nuevas tendencias internacionales en el aprendizaje basado en la tecnología, como parte del movimiento educativo abierto [4]. La popularidad de los MOOCs se explica a partir del momento en el que un conjunto de universidades en el mundo adoptaron este concepto adhiriéndose a diferentes iniciativas, entre ellas edX, Coursera o Udacity [5].

Sin embargo, hay que mencionar que el diseño y construcción de un MOOC presenta algunas complejidades debido a características propias, que muestran divergencias en comparación a los cursos tradicionales. Al crear un curso on-line, no basta con digitalizar los contenidos, y volcarlos en una plataforma, ya que los potenciales alumnos tienen diferentes necesidades y motivaciones, que los conducen por distintas trayectorias en el proceso de aprendizaje [6]. Esto le exige a los docentes habilidades tecnológicas, didácticas y tutoriales que se deben poner en juego al momento de diseñar el curso en forma integral. Una de las principales diferencias respecto a los cursos tradicionales, ya sea presenciales o en línea, reside en el número masivo de estudiantes que puede cubrir un MOOC; por lo que se debe integrar variados recursos que promuevan mayores niveles de compromiso, motivación y aprendizaje en los estudiantes, evitando de este modo, la gran cantidad de alumnos que abandonan el curso una vez comenzado. Por otro lado, los MOOCs han contribuido al replanteo sobre la necesidad de una redefinición de los actuales roles docentes y ha situado al estudiante en un papel de mayor protagonismo en el proceso de selección y autoevaluación de sus aprendizajes, orientándose en estos nuevos modelos de gratuidad y masividad. Se requieren de cambios metodológicos, diseños colaborativos e interactivos, materiales ubicuos y atractivos que faciliten y promuevan la navegación y el descubrimiento, en entornos diseñados para tal fin.

Al comenzar a trabajar con MOOCs es importante establecer una metodología que permita elaborar un marco conceptual que sirva de guía para el diseño del curso desde cero. Los puntos principales a tener en cuenta al momento de definir una

metodología están relacionados con el establecimiento de los objetivos, la planificación del contenido y la definición de las actividades y de las evaluaciones. Además, resulta importante definir la clase de MOOC a crear, en relación a los criterios pedagógicos y tecnológicos, teniendo en cuenta que se pueden clasificar según los modelos a seguir [7].

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Las tareas de investigación, desarrollo e innovación que se vienen llevando a cabo comprenden distintas líneas de trabajo relacionadas a e-learning y cuestiones de interoperabilidad e integración de plataformas de código abierto utilizadas en este campo. En este sentido se compararon distintas herramientas open source para soporte de MOOCs como OpenEdx [8] y openMOOC [9] y se optó por el primero de ellos por las facilidades de instalación y configuración, los módulos disponibles y la amplia comunidad de usuarios que contribuyen al proyecto. La plataforma utilizada es OpenEdX, dado que es una de las más populares en el mercado, actualmente es el segundo proveedor en el mundo en cantidad de cursos ofrecidos. Edx provee toda la funcionalidad básica de las herramientas de MOOCs, en lo relativo a la creación de contenido, gestión de usuarios, confección de evaluaciones y utilización de foros y wikis. En <https://actividades.linti.unlp.edu.ar> se encuentran la plataforma en producción, con un curso implementado en modalidad MOOC, dictado de septiembre a diciembre del 2017.

Asimismo, se aplican los trabajos de investigación que se vienen realizando sobre el tema de Accesibilidad Web,

estudio de las normas de accesibilidad y su aplicación en el desarrollo de sistemas y sitios Web. OpenEdx cuenta con una línea de trabajo específica sobre la temática [10].

Continuando con las líneas planteadas en WICC 2017 [11], en esta etapa el foco de estudio se relaciona con la evaluación del caso de uso desarrollado teniendo en cuenta diferentes criterios, con el objetivos de establecer un punto inicial en la tarea de diseñar y crear cursos masivos. Los criterios a tener en cuenta, incluyen desde la organización del contenido, la interfaz del usuario, herramientas sociales hasta analíticas de aprendizaje y estrategias de evaluación [12]. Se realizará un análisis global del curso y de la plataforma utilizada, en forma interrelacionada, en base a diferentes metodologías de diseño [6] [13].

Respecto del curso se considerarán sus características propias de contenido y lo relacionado a la secuencia de aprendizaje. En este punto, se analizarán en detalle, tanto la organización del curso en las unidades planteadas, como el formato o tipo de contenido incluido en cada una de ellas. Cabe destacar que el curso implementado de accesibilidad Web viene dictándose en modalidad no presencial utilizando la plataforma Moodle, (disponible en cursos.linti.unlp.edu.ar), desde el año 2013. Inicialmente pensado para desarrolladores Web, luego se abrió a toda la comunidad interesada en la temática con un módulo especial para los desarrolladores. Cuenta con tutores formados y colaboradores que se suman año a año para participar de las nuevas ediciones.

Se analizarán en detalle las actividades planteadas, teniendo en cuenta que deben estimular su capacidad de relacionar directamente el contenido con su contexto personal, social o laboral, creando casos concretos de aplicación. El uso de

espacios de comunicación entre los alumnos representa una de las actividades más destacadas y es importante para fomentar la interacción y la generación de conocimiento compartido, logrando una forma de aprendizaje colaborativo. Esta comunicación puede ser centralizada, a través de foros de debate de la propia plataforma, o distribuida en diferentes herramientas de comunicación. Las actividades dentro del curso son necesarias para regular en forma continua el aprendizaje de los estudiantes. Las evaluaciones representan un proceso de regulación de la enseñanza y el aprendizaje y resulta esencial para saber si el alumno ha adquirido las habilidades esperadas. En este punto se analizarán las tareas incluidas en cada unidad, en cuanto a la comprensión de las consignas, el grado de cumplimiento en tiempo y forma por parte de los alumnos y los resultados obtenidos, tanto en las evaluaciones por pares como en las tradicionales corregidas por el profesor.

El estudio realizado será en base a guías metodológicas utilizadas en la Universidad Autónoma y en la Universidad Carlos III, ambas de Madrid. La primera se unió al consorcio edX para difundir a través de esta plataforma sus MOOCs [14].

Respecto a las características de la plataforma, se evaluarán cuestiones de usabilidad y facilidad de uso en cuanto a la localización y acceso a los contenidos; aspectos de instalación y configuración y cuestiones de accesibilidad.

A partir de la experiencia llevada a cabo con la implementación del curso, y la evaluación de los distintos aspectos relacionados con la organización, el desarrollo y los resultados, se analizarán distintas metodologías y herramientas para transformar cursos en línea en MOOCs, analizando y evaluando también nuevas plataformas de base, nuevos

medios de comunicación, contenidos y formatos. El trabajo interdisciplinario con diseñadores visuales, expertos en comunicación, contenidos y educadores permitirá generar una experiencia y un conocimiento enriquecedor, extensible a otras temáticas y líneas de investigación del LINTI.

3. RESULTADOS Y OBJETIVOS

Según las líneas de trabajo descriptas, se plantean los siguientes objetivos:

- Analizar metodologías de planificación y diseño de MOOCs en distintas universidades del mundo.
- Analizar modelos pedagógicos involucrados en la creación de cursos masivos, para evaluar su adecuación a los potenciales cursos.
- Evaluar distintas metodologías y herramientas para transformar cursos existentes en MOOCs, analizando y evaluando también nuevas plataformas de base, nuevos medios de comunicación, contenidos, formatos y evaluaciones.
- Evaluar el MOOC creado desde distintas perspectivas docente, institucional y técnica, a partir del estudio de distintas experiencias.
- Aplicar técnicas de diseño estudiadas a partir del caso de uso desarrollado.
- Como base de la planificación estipulada, analizar los resultados obtenidos en cada etapa, corrigiendo posibles falencias en un ciclo de mejora continua.
- Realizar estudios comparativos entre las dos experiencias que permitan aportar nuevos

conocimientos significativos a los nuevos entornos de aprendizaje y enseñanza que se generan con estas herramientas.

- Medir el grado de aceptación del curso por parte de los participantes y realizar análisis de datos considerando distintas variables como formación, edad, género, procedencia, participación y rendimiento, entre otros.
- Establecer pautas de diseño y construcción de MOOCs que se utilicen como punto de partida para la la creación de cursos masivos MOOCs sobre temas que se investigan y sobre los cuales se viene trabajando en el LINTI.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo se encuentra formado por tres profesoras de amplia trayectoria en el campo de la investigación, que trabajan en el área de ambientes virtuales de aprendizaje y accesibilidad web. Además, dos alumnos de la carrera de Licenciatura en Informática de la Facultad de Informática.

La participación en eventos de la especialidad, ha permitido seguir estableciendo canales de comunicación con otros investigadores que trabajan en las mismas áreas.

5. REFERENCIAS

- [1] Dictamen Proyectos de Extensión UNLP 2015.
<https://unlp.edu.ar/frontend/media/30/3130/9bd20c385bdaa2ffe741ec4cdd418d79.pdf>
- [2] Dictamen Proyectos de Extensión UNLP 2016 - 2017.

<https://unlp.edu.ar/frontend/media/31/3131/fd4f9136c8ad47a0bf47240ae2f65be9.pdf>

[3] Daniel, John. "Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility.", 2012. Disponible en:

<http://blog4222.blogspot.com.ar/2012/09/making-sense-of-moocs-musings-in-maze.html>.

[4] Ramírez, M., Burgos, J. "Movimiento Educativo Abierto. Acceso, colaboración y movilización de recursos educativos abiertos." México: CIITE., 2012

[5] Torres Mancera D., Gago Saldaña, D. "Los MOOCs y su papel en la creación de comunidades de aprendizaje y participación". RIED, Revista Iberoamericana de Educación a Distancia, vol. 17, núm. 1, 2014, pp. 13-34, ISSN: 1138-2783.

[6] Montoro G., Muruzábal O., Sandoval G., Wee C. "7 pasos para diseñar un MOOC de calidad: Una propuesta para la colaboración entre profesores y diseñadores de aprendizaje". Actas de la Jornada de MOOCs en español en EMOOCs 2017 (EMOOCs-ES)

[7] García-Peñalvo, F., Fidalgo-Blanco, A., Sein-Echaluce, M. L. "An adaptive hybrid MOOC model: Disrupting the MOOC concept in higher education". Telematics and Informatics, 2017, <http://dx.doi.org/10.1016/j.tele.2017.09.012>

[8] Open edX Portal | Open Source MOOC Platform <https://open.edx.org>

[9] Open MOOC <https://www.openeducationeuropa.eu/es/institution/open-mooc>

[10] Accessibility Features | Open edX Portal | Open Source MOOC Platform
<https://open.edx.org/features-roadmap/accessibility/all>

[11] Díaz, J., Schiavoni, A., Amadeo, P. "Construyendo un Curso Masivo en Línea sobre Accesibilidad Web Planteos y Desafíos de los Nuevos Entornos de Aprendizaje". WICC 2017, XIX Workshop de Investigadores en Ciencias de la Computación, ISBN: 978-987-42-5143-5, Ciudad Autónoma de Buenos Aires, 27-28 Abril, 2017.

[12] Yousef, A. M. F., Chatti, M. A., Schroeder, U., Wosnitza, M. "What Drives a Successful MOOC? An Empirical Examination of Criteria to Assure Design Quality of MOOCs," *2014 IEEE 14th International Conference on Advanced Learning Technologies*, Athens, 2014, pp. 44-48. doi: 10.1109/ICALT.2014.23

[13] "Guía metodológica para la planificación, diseño e impartición de MOOCs (Massive Open Online Courses) and SPOCs (Small Private Online Courses)". Unidad de Tecnología Educativa e Innovación Docente (UTEID), Universidad Carlos III de Madrid, Versión 3, 15 de octubre de 2014. www.uc3m.es/uteid.

[14] Unidad de Tecnologías para la Educación de la Universidad Autónoma de Madrid.
https://www.uam.es/ss/Satellite/es/1242677588563/subHomeServicio/MOOCs_de_la_UAM_en_edX.htm

ANÁLISIS Y DESARROLLO DE UNA HERRAMIENTA PARA EL MANEJO DE LOS OPERADORES DEL ÁLGEBRA RELACIONAL

Lobos Anfuso, D.; Bustos Aguiar, M. S.; Baquinzay, M. R.; Acosta, D. G.; Palliotto, D.

Departamento de Informática, Facultad de Tecnología y Ciencias Aplicadas,
Universidad Nacional de Catamarca
danielalobosanfuso, solebustosaguiar, diana.palliotto{@gmail.com}

RESUMEN

El estudio de las bases de datos relacionales forma parte de la currícula de las carreras de Informática. Más precisamente, los lenguajes de consulta constituyen una parte importante de estos estudios. Actualmente, en el espacio curricular "Base de Datos" de la carrera de Ingeniería en Informática de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca, se enfrenta el problema de no contar con una herramienta precisa y multiplataforma para realizar consultas aplicando los operadores del Álgebra Relacional. Por este motivo surgió la inquietud de desarrollar una nueva herramienta de software educativo que sea capaz de solucionar el problema mencionado. Se continúa con el trabajo "Análisis comparativo de herramientas para la enseñanza del Álgebra Relacional" presentado en el "Primer Congreso Internacional de Educación en Ciencia y Tecnología", realizado en la Facultad de Ciencias Exactas y Naturales (UNCa), actualizando los resultados obtenidos y realizando nuevos estudios comparativos.

Palabras clave: base de datos relacional, álgebra relacional, operadores relacionales, software educativo.

CONTEXTO

Este trabajo presenta la línea de investigación y desarrollo que se lleva a cabo dentro del

proyecto denominado "Análisis y desarrollo de una herramienta para el manejo de los operadores del Álgebra Relacional" (Código 02/M736), aprobado y financiado por el Consejo de Investigación de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Catamarca (SECyT – UNCa) para el período 2017-2018.

1. INTRODUCCIÓN

En la construcción de Sistemas de Información, el Modelo Relacional, que tiene sólidos fundamentos matemáticos y se basa en la Teoría de Conjuntos, se ha establecido como el principal modelo de datos.

Esta base teórica contribuye al diseño de las Bases de Datos Relacionales (BDR) y, especialmente, al procesamiento eficiente de las consultas. El modelo define de forma precisa los diferentes lenguajes abstractos, siendo uno de estos lenguajes el Álgebra Relacional (en adelante AR). El AR permite a los interesados en el estudio de las BDR, tener un primer acercamiento a los datos y poder manipularlos mediante operaciones sencillas para obtener determinada información.

El espacio curricular "Base de Datos" de la carrera Ingeniería en Informática (FTyCA, UNCa) incluye en su contenido el AR. Durante algún tiempo se buscaron diferentes opciones para que los alumnos puedan aplicar los conocimientos adquiridos en el tema. Esta búsqueda ha resultado inquietante debido,

justamente, a la diversidad de herramientas con las que se puede trabajar, aunque no todas cumplen con los requerimientos de los docentes.

Si bien este tema no presenta ante los alumnos dificultad para entender de qué se trata, sí existe el problema al momento de querer aplicar los operadores mediante una herramienta. Los alumnos desarrollan sus trabajos prácticos y desean verificar los resultados de las operaciones del AR que plantean. Los años de experiencia docente en la enseñanza del AR han demostrado que es complicado para los estudiantes conocer o determinar si son correctas las consultas plasmadas en un papel, en términos del lenguaje relacional, pero las herramientas de software actuales presentan ciertas limitaciones

Existen muchas de estas herramientas pero, en la mayoría de los casos, no permiten el control óptimo de la manipulación de los operadores del AR generando inconvenientes, no sólo para los alumnos, sino también para los docentes involucrados en el estudio de las BDR, impidiendo así un buen desarrollo del proceso de aprendizaje.

En la actualidad, no existe un software multiplataforma con una interfaz intuitiva, fácil y sobre todo una herramienta libre, con un analizador sintáctico que tome como entrada las consultas mediante operadores, las analice, verifique si se trata de una consulta válida y genere nuevas relaciones.

Existen diferentes trabajos comparativos de estas herramientas que mencionan las ventajas y desventajas entre ellas, sus características y sus funcionalidades.

Los docentes de la asignatura "Base de Datos", integrantes de este proyecto, realizaron el trabajo titulado "Análisis comparativo de herramientas para la enseñanza del Álgebra Relacional" que se presentó en el "Primer Congreso Internacional de Educación en Ciencia y Tecnología - Facultad de Ciencias Exactas y Naturales – UNCa". Se pretendía conocer cuál es la herramienta existente más adecuada para trabajar con los operadores del AR. Luego de este análisis, se llegó a la conclusión que las

herramientas estudiadas tenían más desventajas que ventajas.

Con este panorama surgió la motivación para desarrollar una nueva herramienta de software multiplataforma, puesto que el equipo de trabajo cuenta con los conocimientos necesarios para ello.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El espacio curricular "Base de Datos" actualmente carece de una herramienta integral que contemple todos los operadores del AR y que se adapte a todos los sistemas operativos. A partir de este nuevo software, se pretende disponer de una herramienta que satisfaga todas las necesidades y requerimientos en cuanto a los operadores del AR.

De esta manera alumnos y docentes se verán beneficiados. Por un lado, los docentes podrán demostrar que todos los operadores del AR se pueden aplicar y verificar, obteniendo resultados adecuados. Y, desde el punto de vista de los alumnos, podrán reforzar sus conocimientos al mismo tiempo que realicen sus aportes para mejorar el software.

Con este proyecto también se pretende dar a conocer la herramienta en el ámbito educativo y a cualquier interesado en el estudio del AR, contribuyendo a la formación en el tema.

3. RESULTADOS Y OBJETIVOS

El objetivo general que se persigue es el siguiente:

- Obtener una herramienta ágil y multiplataforma para los operadores del AR.

Los objetivos específicos son:

- Realizar un estudio comparativo de las herramientas existentes determinando sus ventajas y desventajas.
- Analizar distintos aspectos, entre ellos los pedagógicos, necesarios para el desarrollo de una nueva herramienta.
- Diseñar e implementar la nueva herramienta.

Actualmente ya se cuenta con un primer prototipo que se utilizó en la asignatura durante el año 2017.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de este proyecto está compuesto por un docente investigador formado, tres docentes investigadores en formación y un estudiante avanzado de la carrera de Ingeniería en Informática. Se proyecta que el estudiante desarrolle su trabajo final de grado dentro de este proyecto.

5. BIBLIOGRAFÍA

Aho, A. V., Sethi, R., & Ullman, J. D. (2000). *Compiladores: principios, técnicas y herramientas*. Pearson Educación.

Chavarría-Báez, L., & Rojas, N. O. (2016). Sobre el uso de herramientas CASE para la enseñanza de bases de datos. *Sistemas, Cibernética e Informática*, 13(2), 51-56.

Date, C. J. (2015). *Introducción a los Sistemas de Bases de Datos*. Argentina: Pearson.

Gardikiotis, S. K., & Malevris, N. (2006). Program Analysis and Testing of Database Applications. 5th IEEE/ACIS International Conference on Computer and Information Science and 1st IEEE/ACIS International Workshop on Component-Based Software Engineering, Software Architecture and Reuse (ICIS-COMSAR'06)

Kinaipp, R. V., & Alves, A. G. (2015). Instrumento Pedagógico para Apoio a Aprendizagem de Álgebra Relacional. *UNOPAR Científica Ciências Exatas e Tecnológicas*, 9(1).

Koloniari, G. (2017, September). Evaluating the Use of an Interactive Software Tool for Learning BCNF Normalization. In *Proceedings*

of the 8th Balkan Conference in Informatics (p. 27). ACM.

Sandoval-Bringas, A., Carreño-León, M., Leyva-Carrillo, A., & Estrada-Cota, I. (2017, October). Experience of a didactic tool as a support in the learning of relational algebra. In *Learning Technologies (LACLO), 2017 Twelfth Latin American Conference on* (pp. 1-4). IEEE.

Seibert, J., Silva, V., Orleans, L., & Zamith, M. (2016, November). CUDA-RA: Uma ferramenta de interpretac, ao de algebra' relacional e estrutura de dados para GPU. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)* (Vol. 27, No. 1, p. 1290).

Silberschatz, A., Korth, H. F., Sudarshan, S., Pérez, F. S., Santiago, A. I., & Sánchez, A. V. (2002). *Fundamentos de bases de datos*. McGraw-Hill.

Xohua-Chacón, A., Benítez-Guerrero, E., & Mezura-Godoy, C. (2017, September). TanQuery: a tangible system for relational algebra learning. *Interacción '17: Proceedings of the XVIII International Conference on Human Computer Interaction*.

EVALUACIÓN DE CALIDAD DE HERRAMIENTAS UTILIZADAS EN LA ENSEÑANZA DE LA PROGRAMACIÓN BASADA EN ISO 25000

María C. Espíndola, Cristina L. Greiner, Gladys N. Dapozo,

Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura

Universidad Nacional del Nordeste. 9 de Julio N° 1449. Corrientes. Argentina.

{gndapozo, cgreiner}@exa.unne.edu.ar, mcespindola@yahoo.com

RESUMEN

En la actualidad la enseñanza de la programación ha adquirido una especial importancia. El Consejo Federal de Educación declaró de importancia estratégica para el sistema educativo argentino la enseñanza y el aprendizaje de la programación durante la escolaridad obligatoria, para fortalecer el desarrollo económico-social de la Nación. La demanda insatisfecha de la industria del software promueve la formación de perfiles profesionales para el desarrollo de software. Por otra parte, se sostiene que la programación fomenta habilidades vinculadas con el pensamiento computacional, habilidades requeridas en el siglo XXI para entender el mundo y la tecnología que nos rodea. En este proyecto se plantea el desarrollo de un marco de referencia para evaluar las características de las herramientas visuales o lúdicas utilizadas en la enseñanza inicial de la programación, y permita orientar a los docentes hacia la selección de las herramientas más adecuadas para incorporar en sus propuestas docentes, en función de los objetivos específicos de aprendizaje, nivel educativo de los niños o jóvenes y expectativas del docente.

Palabras clave: Enseñanza de la programación inicial. Herramientas visuales o lúdicas. Evaluación de calidad de software.

CONTEXTO

Esta propuesta es parte del plan de Trabajo Final de la Maestría en Tecnologías de la Información (UNNE), que se desarrolla en el marco del proyecto 16F016 “Promoción del Pensamiento Computacional para favorecer la formación en STEM”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE).

1. INTRODUCCIÓN

El crecimiento de la industria del software en nuestro país no es una novedad. “El sector industrial de software es uno de los que más ha crecido en Argentina durante la última década, e incluso actualmente Argentina ya exporta más software que carne” [1] El último reporte de la Cámara Empresarial de Software y Servicios Informáticos (CESSI), en función de datos del 2016 y expectativas del 2017, señala que en 2016 la industria del software volvió a generar empleo, dejando una demanda insatisfecha de más de 5.000 puestos sin cubrir [2].

Necesidad de Recursos Humanos

Por otra parte, un estudio reciente señala que es evidente la producción limitada de graduados en STEM respecto de las necesidades del campo productivo y, la falta de masa crítica en capacidades para trabajar en la innovación digital [3].

Consciente de esta situación, diversas políticas públicas en nuestro país promueven la industria del software. Sin

embargo, éstas no solo apuntan al desarrollo del software, sino que van más allá, porque en la actualidad el mundo que nos rodea depende cada vez más del uso de la tecnología, que alcanza a las actividades profesionales, de educación, la comunicación en el trabajo y con los vínculos afectivos, el tiempo de esparcimiento, entre otras, atravesadas por el uso de dispositivos tecnológicos, de los cuales pocos saben cómo funcionan, y del trabajo que realizan las personas que se desarrollan en ese campo. Es así que aprender a programar se ha transformado en los últimos años en un componente clave de las “competencias del Siglo XXI” [4]. Permite desarrollar habilidades de pensamiento que sirven para cualquier tipo de actividad, como la capacidad de abstracción y de planificación, la descomposición de problemas, el trabajo en equipo, entre otras.

Enseñanza de la programación

En apoyo a esta iniciativa, el Consejo Federal de Educación ha declarado de importancia estratégica para el sistema educativo argentino la enseñanza de la programación durante la escolaridad obligatoria, para fortalecer el desarrollo económico-social de la Nación.

Por lo tanto, la enseñanza de la programación es una problemática que atraviesa todos los niveles [5], [6], [7]. Sin embargo, no es una tarea trivial. La investigación existente plantea muchas discusiones e ideas sobre la mejor manera de enseñar la programación introductoria, y las dificultades que enfrentan los estudiantes. Entre estas, el hecho de que los estudiantes consideran a la programación como una actividad puramente técnica en lugar de un conjunto de habilidades combinadas de resolución de problemas. Tienden a desarrollar un conocimiento superficial y no crean estrategias de resolución de problemas a través de la programación [8]. Por este motivo se implementan diferentes estrategias, entre las que

sobresale la de enseñar con herramientas lúdicas [9], [10].

En este contexto, cobra relevancia la elección de la herramienta a utilizar. En esta línea de investigación se propone la elaboración de un marco de referencia que brinde una guía a la hora de planificar un curso de enseñanza de programación introductoria.

Herramientas lúdicas

Los NPEs (Novice Programming Environments - Entornos de programación para principiantes) han tenido amplio reconocimiento en los últimos años, ya que la investigación ha demostrado que desempeñan un papel importante en la motivación de los nuevos programadores. En lugar de codificar instrucciones, un NPE ofrece características visuales que ocultan la complejidad de la sintaxis del lenguaje de programación, facilitando a los principiantes el uso de las estructuras algorítmicas básicas [11].

En [12] los autores presentan un modelo para desarrollar el pensamiento computacional basado en la programación visual. Este concepto se basa en la idea What You See Is What You Get (WYSIWYG), que evita a los programadores la escritura de código, ya que pueden usar bloques visuales para construir funcionalidades.

Mediante una revisión de la literatura, se identificaron cinco factores de éxito para enseñar programación con un enfoque lúdico: la motivación del estudiante, la integración y la participación en clase, el enfoque centrado en el alumno, la interacción y retroalimentación, y la fluida integración del contenido educativo en el juego [13].

Sin embargo, en [14] los autores señalan que, si la herramienta utilizada y los objetivos de aprendizaje no están alineados, podría no lograrse un gran efecto en el aprendizaje deseado. Señalan también la falta de certeza en la elección de la mejor herramienta para el

logro de los objetivos de aprendizaje, basada en los atributos de la misma.

Marcos de referencia (MR)

La utilización de un MR [15]:

- Evita tener que crear indicadores, ya que están definidos en el modelo.
- Permite disponer de un marco conceptual completo.
- Proporciona objetivos y estándares iguales para todos, en muchos casos ampliamente contrastados.
- Posibilita medir con los mismos criterios a lo largo del tiempo, por lo que es fácil detectar si se está avanzado en la dirección adecuada.

Algunos marcos teóricos tienden a darle mucho énfasis a los aspectos técnicos del software educativo. Sin embargo, un diseño técnico sólido no garantiza el valor educativo. No solo se debe medir el software educativo como producto, sino también los procesos de enseñanza-aprendizaje que facilita o potencia [16].

En este trabajo se propone construir un MR que permita realizar una evaluación integral de herramientas lúdicas. Se considera evaluar la calidad del producto mediante un estándar internacional. Estos permiten la definición de criterios de evaluación, la especificación de requerimientos, la descripción de componentes en relación a ellos y la identificación de desajustes de manera sistemática facilitando el proceso de evaluación y selección del software [17].

Estándar ISO/IEC 25010

Permite identificar las características de calidad relevantes para establecer criterios de satisfacción y las medidas correspondientes. Está compuesto por ocho características y subcaracterísticas de calidad. Para este trabajo se propone el análisis y selección de las características y subcaracterísticas que resulten relevantes para una evaluación sistemática y práctica de las herramientas lúdicas para la enseñanza de la programación.

Existen numerosos trabajos sobre evaluación de calidad de dispositivos con fines educativos [16], [18], [19].

Todos los autores coinciden en que a la hora de evaluar estos artefactos es necesario considerar dos aspectos:

- calidad desde el punto de vista técnico
- calidad en cuanto al impacto en el aprendizaje.

En [20] se realiza un análisis de varios modelos en los que surgen dos categorías importantes en la evaluación de software educativo: la dimensión pedagógica y la técnica. La parte pedagógica contempla las características de los destinatarios, los objetivos de enseñanza, los contenidos y el aprendizaje que promueve la aplicación. Desde el punto de vista técnico se resaltan la usabilidad y las características de la interface.

Criterios para determinar el impacto en el aprendizaje

En [21] se analizan trabajos relacionados al uso de los JS, y los autores concluyen que su uso incrementa significativamente el aprendizaje en el área de la Informática. Señalan también que en los artículos que se revisaron, el aspecto que más se evaluó además del aprendizaje fue la motivación. Esto se debe a que la motivación es considerada como un factor importante que describe lo que hace un buen juego independientemente de sus cualidades educativas [22]. En una experiencia basada en el desarrollo de actividades lúdicas presentada en [9] se concluyó que esta metodología permitió despertar el interés de los alumnos y consolidar conceptos básicos de programación.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta línea, se proponen las siguientes actividades:

- Profundizar el conocimiento sobre marcos metodológicos que sustentan la evaluación de software, desde el

punto de vista de la ingeniería del software y desde los procesos de enseñanza-aprendizaje que potencia.

- Relevar y evaluar herramientas utilizadas en la enseñanza inicial de la programación, y realizar una clasificación según sus características
- Definir criterios de evaluación de las herramientas, en función de los distintos aspectos que surgirán de la etapa de relevamiento y evaluación.
- Diseñar un marco de referencia que oficie de guía en la selección de la herramienta más apropiada para el logro de los objetivos de aprendizaje.
- Validar el marco de referencia en una situación concreta, a definir de acuerdo a las posibilidades de acceso a los distintos ámbitos educativos.

3. RESULTADOS ESPERADOS

Se espera obtener un marco de referencia que oriente la selección de las herramientas visuales o lúdicas orientadas a la enseñanza inicial de la programación.

Este aporte facilitará a los docentes de los niveles educativos no universitarios, contar con información que les orientará hacia la selección de las herramientas más adecuadas para incorporar en sus propuestas docentes, en función de los objetivos específicos de aprendizaje, nivel educativo de los niños o jóvenes y expectativas del docente.

Este recurso contribuirá a facilitar el logro de los objetivos de las políticas públicas que fomentan la incorporación de la programación en la escuela, como una competencia necesaria para la formación del siglo XXI.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea participan dos investigadoras formadas y una tesista de posgrado que realiza su trabajo final para obtener el título de Magister en Tecnologías de la Información, oferta interinstitucional entre la UNNE y

UNAM, dictado 2016-2017.

5. REFERENCIAS

- [1] SW 48 - SG Software Guru #48 Published on Sep 22, 2015, Revista SG. Software Guru. Mexico, Mexico. ISSN: 1870-0888
- [2] Reporte anual sobre el Sector de Software y Servicios Informáticos de la República Argentina. (OPSSI – CESSI). Mayo 2017.
- [3] Katz, R. (2016). TIC, digitalización y políticas públicas. En Entornos Digitales y Políticas Educativas. IPE-UNESCO.
- [4] Borchardt, M.; Roggi, I. (2017) Ciencias de la Computación en los Sistemas Educativos de América Latina. SITEAL. <http://www.siteal.iipe-oei.org>
- [5] Dapozo, G.; Greiner, C.; Pedrozzo Petrazzini, G.; Chiapello, J. Despertando vocaciones TIC mediante juegos y animaciones. Revista: Extensionismo, Innovación y Transferencia Tecnológica. Claves para el Desarrollo. ISSN 2422-6424 Vol. 1 Año 2014. Pag. 80-89.
- [6] Dapozo, G.; Greiner, C.; Petris, R. “Enseñar programación desde los juegos a lo formal”. 1er Congreso de Educación y Tecnologías del Mercosur: de la digitalización a la virtualización. ISBN 978-987-3619-26-7. 2017.
- [7] Dapozo, G.; Petris, R.; Greiner, C. “Programación en las escuelas. Experiencia de formación docente en el Nordeste Argentino”. REMEIED: Memorias del Encuentro Internacional de Educación a Distancia. Desafíos de la cultura digital en América Latina. ISSN: 2395-8901. Diciembre 2016.
- [8] C. Kazimoglu, M. Kiernan, L. Bacon y L. Mackinnon, “A serious game for developing computational thinking and learning introductory computer programming,” *Procedia - Social and Behavioral Sciences* 47, 2012.
- [9] Dapozo, G.; Greiner, C.; Petris, R.; Espíndola, M.C.; Company, A. Introduction To Programming Based On Playful Activities In The University. XLIII Latin American Computer

- Conference (CLEI) IEEE Xplore digital library. December 2017.
- [10] Dapozo, G.; Greiner, C.; Petris, R.; Espíndola, M.C.; Company, A. Enseñanza de la Programación en la Universidad. Factores que inciden en el buen desempeño de los estudiantes. Actas 5to Congreso Nacional de Ingeniería Informática / Sistemas de Información (CONAIISI 2017). ISSN: 2347-0372. Noviembre de 2017.
- [11] J.C. Olabe, M.A. Olabe, X. Basogain y C. Castaño, "Programming and robotics with Scratch in primary education," *Education in a Technological World: Communicating current and Emerging Research and Technological Efforts*, Badajoz - Spain: Formatex, 356–363, 2011.
- [12] Min Xiao, Xiaohua Yu. A Model of Cultivating Computational Thinking Based on Visual Programming. The Sixth International Conference of Educational Innovation through Technology. 2017.
- [13] R. Heininger, L. Prifti, V. Seifert, M. Utesch and H. Krcmar, "Teaching how to program with a playful approach: A review of success factors," 2017 IEEE Global Engineering Education Conference (EDUCON), Athens, 2017.
- [14] Daisuke Saito, Ayana Sasaki, Hironori Washizaki, Yoshiaki Fukazawa, Yusuke Muto. Program Learning for Beginners: Survey and Taxonomy of Programming Learning Tools. 2017 IEEE 9th International Conference on Engineering Education (ICEED), Kanazawa, Japan.2017.
- [15] García de la Torre, M.; Mingo Sarto, L.A.; Sáez de Eguilaz Larreta, M.J.; Arteaga Fernández, L.; Labarta Aizpún, M.; Aragón Marín, L; Ambrosio Flores, P.; Piernavieja Marrón, E. Modelo europeo de excelencia. Adaptación a los centros educativos del modelo de la fundación europea para la gestión de calidad. Ministerio de Educación, Cultura y Deporte
- [16] Carlos Vargas Castillo. Cuatro Modelos para Evaluar el Software Educativo. CACIC 2003.
- [17] Bermeo, J., Sánchez, M., Maldonado, J., & Carvallo, J. (2016). Modelos de Calidad de Software en la Práctica: Mejorando su Construcción con el Soporte de Modelos Conceptuales. Universidad de Cuenca, Ecuador.
- [18] Sánchez, M.A. (2016). Assessing the quality of MOOC using ISO/IEC 25010. 2016 XI Latin American Conference on Learning Objects and Technology (LACLO), 1-4.
- [19] Stella Maris Massa, Patricia Pesado. Evaluación de la usabilidad de un Objeto de Aprendizaje por estudiantes. Revista Iberoamericana de Educación en Tecnología y Tecnología en Educación N°8. ISSN 1850-9959. Diciembre 2012.
- [20] Cova, Ángela, Arrieta, Xiomara, & Reveros, Víctor. (2008). Análisis y comparación de diversos modelos de evaluación de software educativo. *Enlace*, 5(3), 045-067.
- [21] L. García-Mundo, J. Vargas-Enríquez, M. Genero y M. Piattini, "¿Contribuye el Uso de Juegos Serios a Mejorar el Aprendizaje en el Área de la Informática?," *Actas de las XX JENUI*. Oviedo, 2014 ISBN: 978-84-697-0774-6
- [22] T. Connolly, L. Boyle y T. Hainey, "A Survey of Students' Motivations for Playing Computer Games: A Comparative Analysis," *En Proceedings of the 1st European conference on games-based learning (ECGBL)*, Paisley, Scotland, 71-78, 2007.

Tesis Doctorales

Resumen Tesis Doctoral

Título:

Agent SocialMetric: herramienta de asistencia al docente en el dominio Educativo del aula

Autor: Dra. Antonieta Kuz

Director: Dra. Roxana Giandini

Co-Director: Dra. Claudia Pons

Universidad donde se desarrolló la tesis: Universidad Nacional del Centro de la Provincia de Bs As.

Acreditado por Resolución de CONEAU N° 745/11 - Categoría "A". Resoluciones Ministeriales N° 0052/99 y 0665/99.

Fecha de defensa y aprobación: 15/12/2017

RESUMEN

Las Tecnologías de la Información y la Comunicación (TICs) están directamente vinculadas a la naturaleza del aprendizaje, dado que éste se basa en el manejo de información. El uso de las TICs puede enriquecer la interacción entre alumnos y docentes en el contexto de las actividades educativas. Consideramos entonces la necesidad de construir una herramienta tecnológica innovadora de software educativo, que involucre los hábitos de comunicación de los estudiantes y las relaciones entre compañeros para la realización de actividades, permitiendo obtener información tanto del nivel de integración de cada alumno, como del ambiente donde se desarrolla. La herramienta llamada Agent SocialMetric, desarrollada específicamente para esta investigación integra las líneas provenientes de los Agentes de Interfaz Conversacionales y del Análisis de Redes Sociales (ARS). Para comprender y abordar la tarea de investigación, la metodología que consideramos se fundamenta en tres grandes fases: la investigación y desarrollo inicial, la investigación aplicada, y la transferencia. En base a las dos líneas tecnológicas mencionadas, integradas a la metodología de investigación, se provee un marco de trabajo innovador, desarrollado por etapas en un proceso de trabajo Ágil. En este marco se busca contribuir a las

Ciencias de la Computación, con un recurso tecnológico educativo innovador que le permita al docente superar limitaciones existentes en la apropiación de las TICs. Como resultado, dicho recurso ofrece la asistencia al docente facilitándole el diagnóstico en el contexto del clima de aula, sobre la dinámica y el funcionamiento de la clase. También brinda al docente la posibilidad de conocer e interpretar los vínculos que se generan entre los alumnos, y de abordar una intervención más precisa y ajustada a las interacciones sociales. Además el proyecto incluye el análisis de varios casos de estudio como apoyo a la práctica docente, y de usabilidad del software desarrollado. Finalmente se contrasta el software propuesto respecto a otras herramientas existentes en el mercado.

1. MOTIVACIÓN

En la actualidad, encontramos que los docentes deben asumir la responsabilidad de orientar las claves relacionales entre los estudiantes como una vía para prevenir, intervenir y favorecer las relaciones personales para que puedan ser saludables y equilibradas, y así, mejorar el clima de convivencia escolar. A raíz de ello, el docente es el gestor de lo que ocurre en el aula, y por consiguiente en clima que se genere, siendo su papel importante en las relaciones entre alumnos que se realicen dentro de la clase, y en el mejoramiento de las mismas, y en la prevención

de posibles situaciones de conflicto.

Por lo tanto, el profesor debe afrontar dos factores claves: por un lado, conocer las relaciones entre los alumnos, determinando sus posiciones ya sea que encarnen los roles de líderes o de rechazados; y por el otro, abrir un canal de comunicación con los alumnos rechazados y aislados a establecer dicha comunicación. Frente a estos factores es clave que los docentes cuenten con herramientas TICs, que sean confiables, de fácil acceso, que se encuentren disponibles cuando las necesite y que les permita intervenir en el grupo social oportunamente basado en el conocimiento de las relaciones sociales que fluyen en él. El diseño de estrategias didácticas representa un eje integrador de los procedimientos que permiten al estudiante construir sus conocimientos a partir de la información que se suministra en el aula.

Según el estudio elaborado por el PASEM [20] revela que *“la integración de las TIC en la formación docente en la Argentina vinculada a un uso deseable de las TIC en el contexto de la enseñanza, dista de ser unívoca y clara”*. Ellos consideran en su estudio que hay un camino por recorrer en la incorporación de las TICs a la formación docente, sin embargo tiene la pretensión de constituirse en prácticas pedagógicas más complejas y significativas. Asimismo a través de esta investigación observan que cuando los docentes son capaces de asociar las herramientas de software con sus objetivos de la enseñanza, en el marco del proceso de enseñanza-aprendizaje, advierten que el software logra ser un importante movilizador pedagógico que no debe perderse de vista. Además sostienen que *“resulta bastante común la perspectiva de que el desarrollo profesional docente es una tarea y una responsabilidad de cada docente, sujeta a los valores, voluntades y disponibilidades de cada profesor en particular. La integración pedagógica de las TIC en el aula queda en las manos del docente, de cada docente, y del tiempo personal que quiera/pueda dedicar a la propia formación, a la exploración de recursos y/o a la planificación de nuevas secuencias didácticas”*.

A pesar del potencial de las TIC para transformar los entornos educativos, diversos factores

influyen en su bajo nivel de adopción, siendo el ámbito educativo donde la adopción de las TIC ha sido menor o más lento que en otros sectores de desarrollo de la sociedad. Estos factores son conocidos como barreras para la integración de las TIC. Una barrera es definida como *“cualquier condición que haga difícil realizar progresos o lograr un objetivo”* [21]. El conocer y tomar conciencia de las barreras es un aspecto fundamental, para generar las condiciones para la integración de la tecnología. Podemos detallar y esbozar algunas de las barreras más conocidas: la falta de tiempo; problemas técnicos y actitudes de los profesores hacia la tecnología; el acceso a la tecnología; el no contar con el conocimiento suficiente y habilidades con respecto a las TICs [11] [19] entre otras. Según PASEM [20] en estudio que realizan se rescata que existe una barrera de carácter personal en el docente, referida al sentido de auto-eficacia en el uso de las TICs en su aula, a su percepción del esfuerzo que le implica preparar y realizar una clase con TICs y cuán útil estime que será hacer este esfuerzo en términos de mejorar el aprendizaje y una mayor motivación de sus estudiantes. Finalmente ellos concluyen que *“es importante considerar la actitud de los docentes frente a las TIC y medir su percepción que tienen de ellas”*. Además de estas barreras se soslaya la idea de que muchas veces las TICs y en particular el software no son del todo específicos y útiles para resolver los problemas cotidianos que se abordan en el aula, los docentes.

La pregunta permanente de los docentes sobre cuál es la mejor manera de facilitar el clima del aula, ha guiado un numeroso cúmulo de investigaciones y se ha convertido en el eje central de muchas discusiones. Así, los intentos y propuestas metodológicas, son precisamente las que han dado origen a los modelos de enseñanza. El docente es la persona que desempeña el papel más importante en la tarea de ayudar a los estudiantes a adquirir diferentes capacidades, además de ser el responsable de diseñar tanto oportunidades de aprendizaje como el entorno propicio en el aula. Por esto, es fundamental que todos los docentes estén preparados para ofrecer esas oportunidades a sus estudiantes, a través de herramientas que lo posibiliten. Como vemos

aún se carece de un profundo conocimiento acerca de cómo asistir a los profesores en el aula a través de la tecnología y en particular del software, teniendo en cuenta, las relaciones y las interacciones informales que mantienen los alumnos como medio y las barreras tecnológicas que se presentan.

Además de las barreras mencionadas que generan reticencia, una de las fuertes preocupaciones de los docentes es la toma de decisiones en torno a la forma de la problemática del clima del aula. El ambiente o clima del aula es una combinación de variables dentro de un salón de clases que trabajan juntas para promover el aprendizaje en un ambiente cómodo. La Lic. María Antonia Casanova afirma que *“si se quiere actuar de manera social, con métodos sociales y con fines específicos de socialización, es imprescindible conocer a fondo, por parte del profesor, el curso que tiene entre manos, desde un punto de vista sociológico, y utilizar también métodos de este tipo para corregir los posibles desajustes de cualquiera de sus alumnos”* [4].

Teniendo en cuenta esta situación, nos planteamos la necesidad de realizar una investigación que explore mecanismos de obtención y de comunicación efectiva de las relaciones sociales y los hábitos de comunicación en entornos presenciales y virtuales de los estudiantes, evaluada desde la percepción docente. La principal motivación fue trabajar sobre estas barreras que se presentan a través del uso de las TICs y del software educativo, en los docentes dadas por la resistencia por parte de los profesores al manejo de herramientas, y a su potencial didáctico, la escasa integración de las TIC y en particular el software educativo al proceso de enseñanza-aprendizaje, la desatención del clima del aula como factor de repercusión condicionante el proceso de enseñanza y la desestimación de la visión particular del docente.

2.LA TESIS

El objetivo principal es desarrollar una herramienta prototipo de soporte tecnológico web, denominada Agent SocialMetric, de desarrollo propio, para asistir al docente mediante un asistente virtual en el ambiente áulico, soportado en la integración del ARS

(Análisis de Redes Sociales) y Agentes Inteligentes de Interfaces Conversacionales. Se establecerá un modelo abstracto que se concretará bajo un enfoque metodológico basado en una combinación de dos metodologías Thales [14] con Ágile [10], que incluye el desarrollo por etapas de un diseño funcional y arquitectónico, con el fin de crear un asistente virtual interactivo, llamado Albert embebido en la herramienta. Dicha entidad de interfaz gráfica hipermedia se presenta forma de amigable permitiendo que el docente, como pieza esencial en este modelo y en el proceso enseñanza-aprendizaje, obtenga de forma dinámica la estructura interna del aula, que podrá usar para diagnosticar el ambiente áulico e intervenir de acuerdo a los resultados obtenidos.

Dentro del ambiente educativo, existen diversas problemáticas emergentes que requieren un abordaje rápido y efectivo como el bullying, el aislamiento, la repitencia, la falta de integración y adaptación, inclusividad, entre otras; que afectan las interacciones, el clima social y el ambiente de aprendizaje en el aula. A través de Agent SocialMetric será viable identificar y caracterizar los valores sociométricos de los alumnos tanto de forma individual como grupal, posibilitando la comprensión de la estructura del aula mediante la aplicación y cálculo de métricas e indicadores, como los tipos, valores e índices sociométricos.

Este proyecto buscamos como meta, a través de un desarrollo tecnológico integrador e innovador, crear una herramienta de software educativo de autoría propia la cual será capaz de entender y analizar los hábitos de comunicación de los estudiantes y las relaciones entre compañeros, para la realización de tareas, al compartir actividades, permitiendo obtener información tanto del nivel de integración de cada alumno/a como de los contextos en los que se desarrolla. De esta forma la herramienta ofrecerá asistencia al docente para la toma de decisiones respecto a la dinámica y el funcionamiento del clima del aula, contribuyendo así a mejorar la calidad del proceso de enseñanza/aprendizaje.

Hemos plasmado en Agent SocialMetric una integración de dichos conceptos, mediante el desarrollo de un sistema integrador online que devela las percepciones de los estudiantes acerca

de las relaciones interpersonales en un aula de clases, permite el análisis de la interacción existente entre los estudiantes y el docente dentro del marco temático, posibilitándole al docente la toma de decisiones pedagógicas precisas. En la Figura 1, puede visualizarse el modelo de la herramienta integrando el ARS y los agentes inteligentes; con sus distintos niveles de abstracción.

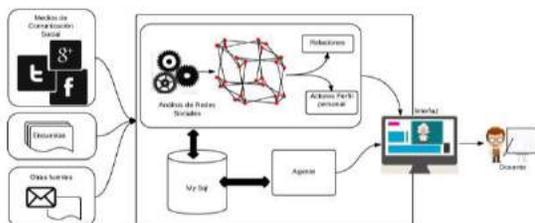


Figura 1. Modelo de la herramienta AgentSocialMetric

El esquema general de nuestro Modelo de Solución puede verse en la Figura 2 que tendrá la versatilidad de permitirle al profesor además, intervenir intencionalmente y oportunamente para prevenir y/o modificar situaciones conflictivas que se puedan presentar en el aula [23] [24].

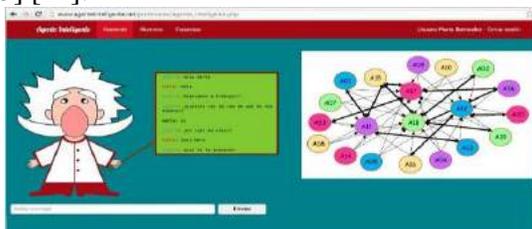


Figura 2. Pantalla principal de Agent SocialMetric

Plasmado en Agent SocialMetric [13] un sistema integrador online que devela las percepciones de los estudiantes acerca de las relaciones interpersonales en un aula de clases, permite el análisis de la interacción existente entre los estudiantes y el docente dentro del marco temático, posibilitándole al docente la toma de decisiones pedagógicas precisas. En la Figura 1 en el Modelo de Solución propuesto, pueden visualizarse los componentes del modelo de nuestra herramienta integradora, que serán esbozados brevemente a continuación, y que posteriormente serán profundizados a lo largo del desarrollo de la tesis. Las fuentes de datos, es una de las partes más importantes del modelo de

la herramienta son las fuentes de datos y la recolección de información. Ahora bien, para crear una red social se debe incluir una lista de todos los actores; pero como los métodos de redes se centran en las relaciones entre los actores, éstos no pueden ser muestreados independientemente.

Al incorporar un Agente conversacional, denominado Albert, el mismo toma un rol activo en la conversación, siendo parte de sus habilidades el reconocer y responder a entradas por parte del usuario así como generar las salidas correspondientes y realizar funciones conversacionales. Es indispensable su aplicación en el aula, debido a que es allí donde los alumnos no sólo aprenden nuevos conocimientos sino que además incorporan y se familiarizan con normas de conducta y diversas formas de desenvolverse ante el sinfín de situaciones que se les presentan. El desarrollo de un bot conversacional es una tarea de alta complejidad. Particularmente, nuestro asistente virtual presenta dos módulos: el de Gestión del Diálogo dentro de la capa de web services y el de Generación de Locución (ver Figura 3).

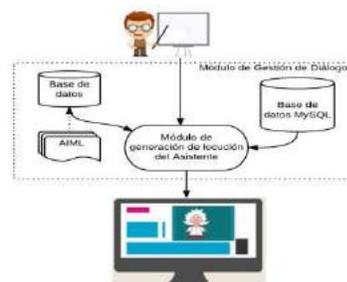


Figura 3. Gestor de Locución de Agent SocialMetric

3. EXPERIMENTOS REALIZADOS

Con el fin de abordar la tarea de investigación, la estrategia que consideraremos está basada en los modelos de Martin y McClure [15]. Consiste en tres grandes fases: la investigación y desarrollo inicial, la investigación aplicada, y la transferencia, que podemos ver en la Figura 4. El tipo de enfoque de investigación seleccionado en nuestro proyecto, es el denominado “desarrollo experimental” [3]. Este enfoque tiene como objetivo aprovechar los conocimientos adquiridos en la investigación y aplicarlos de

forma práctica, obteniendo un resultado que queda plasmado en un prototipo. Encontramos como referencia los trabajos de Basili [3] y Zelkowitz [25] que se centran en el carácter experimental de los proyectos de investigación relacionados a la ingeniería de software.



Figura 4. Esquema de la estrategia de investigación

Centrados en el análisis de casos que presentaremos buscamos validar el conocimiento y la comprensión de los procesos de diagnóstico e intervención llevados a cabo a través de Agent SocialMetric y las técnicas empleadas obtenidas para la validación de los objetivos perseguidos en la creación de la herramienta expuesta. En los casos siguientes realizaremos el análisis del problema, continuando luego con el diagnóstico de las interacciones y las reflexiones acerca de la herramienta [7].

En cada uno de estos casos, preliminarmente se ha explicado a los participantes el propósito y el alcance del sistema. A partir de ahí se pasa a la de experimentación y prueba, especificando el entorno y con qué participantes se llevará a cabo la prueba (además la misma prueba es de carácter individual, a los participantes que favorezca la reflexión y la elaboración personal de los elementos descriptivos principales, para lograr los mejores resultados).

Realizamos el contraste y evaluación, cuya finalidad es el análisis de la situación planteada en cada caso y la búsqueda del sentido que tienen los datos estudiados. También incorporamos pruebas de usabilidad, encuadradas en la experiencia del usuario docente, en función de la actitud del docente ante el software educativo presentado como innovación educativa.

Al llevar nuestra herramienta Agent SocialMetric a un contexto educativo, se requiere que el docente como actor involucrado maneje y considere a las TICs como una metodología

activa de cambio y una preocupación por una formación integral. Este tipo de condiciones permite que los casos se ajusten a los objetivos que busca la herramienta, ser un instrumento que proporcione información valiosa y relevante a toda la comunidad educativa, y que se extienda asimismo su campo de aplicación.

En nuestra investigación hemos determinado tres escenarios áulicos, pertenecientes al nivel primario, secundario y universitario (ver Figura 5) en cada uno de estos escenarios se evaluarán diferentes consignas describiendo las situaciones en función de las problemáticas a abordar. Nos enfocaremos en mostrar tres casos de estudio y su aplicación práctica trabajando tres poblaciones diferentes circunscriptas en las problemáticas dadas en los contextos educativos y escolares.

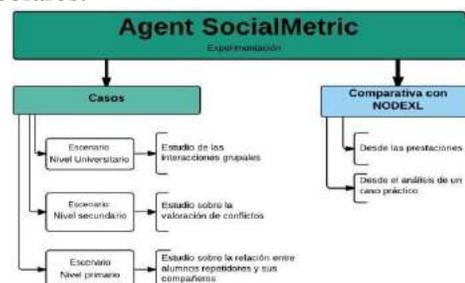


Figura 5. Experimentación a través de Agent SocialMetric

4. APORTES A LA DISCIPLINA

Circunscriptos en las Ciencias de la Computación podemos decir que nuestro aporte principal está dirigido a la Tecnología Educativa, desde un enfoque multidisciplinar [17]. Teniendo presente la definición de Moreira [18], el cual considera a la Tecnología Educativa como un espacio cuyo objeto de estudio son las TICs en cuanto a sus formas de desarrollo, representación, difusión, conocimiento y acceso aplicables en los distintos contextos educativos. En el trabajo de tesis se planteó la construcción de un software educativo [22] de autor como un material o recurso didáctico entendido como un tipo particular de tecnologías de información y comunicación. Este recurso es entonces un medio empleado por el docente para apoyar, complementar el proceso educativo que dirige u orienta. El mismo se basa en la sinergia de varios campos del saber, que convergen en la

generación de un producto computacional: el ARS y los agentes de software conversacionales encuadrados dentro de la ingeniería de software. Las mayores contribuciones de esta tesis son resultado de varios años de investigación sobre estas temáticas y en este camino hemos conseguido aportar a:

- Al **potencial didáctico** del software educativo, a través de la integración de varios campos computacionales, encuadrando esta articulación a la ingeniería de software educativo [9].

- Ampliar el **repertorio de propuestas de software didáctico** destinado a los docentes. En esta obra brindamos una herramienta de apoyo como una propuesta didáctica enriquecedora.

- **Innovar en la implementación práctica del agente conversacional** a través del gestor de diálogos en AIML en idioma español, embebido en Agent SocialMetric donde el problema del diseño y construcción de este gestor va más allá de seguir las ideas y los métodos utilizados en otros procesadores de este tipo, debido a que los existentes cuentan con errores de compilación y no están diseñados para ser implementados en idioma español.

- **Integrar el ARS al campo computacional** a través de la representación y extracción de información, mediante la implementación práctica y modelada en nuestro sistema computacional.

- Combinar metodologías tradicionales con Kanban [16] para obtener **una metodología extendida desde una visión integradora**, justificando la combinación que mejor se vincule y adapte al proceso de desarrollo del software educativo.

- Extraer de las pruebas datos de la interacción e impresiones subjetivas de los usuarios, lo que permite realizar un análisis de los datos obteniendo una serie de resultados y conclusiones acerca de la influencia de la tecnología proporcionada por la herramienta, adaptada a diferentes necesidades.

Por lo tanto, la elección de este tema de tesis reúne desde una perspectiva multidisciplinaria, tres tipos de interés:

a. Pedagógico: dado por el aporte a la inclusión del docente en la utilización didáctica de las

TICs. Actualmente las TICs tienen un papel fundamental en la creación de oportunidades de una manera práctica y útil. Permitiendo que el profesorado no permanezca ajeno a las posibilidades de aprendizaje que de ellas puedan derivarse. Agent SocialMetric brinda una nueva estrategia para el uso de las TICs en la práctica docente, y en el proceso de enseñanza-aprendizaje. Desde una perspectiva metodológica, del estudio de casos derivados del uso de Agent SocialMetric, encontramos por un lado, la detección de problemas de inadaptación social. Esta problemática se detecta en la existencia de elementos aislados en el grupo, no elegidos por nadie, ante lo cual el docente está advertido de esta anomalía para procurar que los demás les presten atención. También incluye la detección de alumnos rechazados por la mayoría de los miembros y la detección de subgrupos y los líderes. Por otro lado, la determinación de estrategias metodológicas en el aula. Esto se da especialmente en los momentos previstos para el trabajo en grupos operativos a lo largo del desarrollo de una unidad didáctica. La herramienta brinda la estructura social o informal del grupo en su conjunto. Se procurará, en principio, que los grupos de trabajo se ajusten a los grupos formados espontáneamente, en función de las preferencias manifestadas por el alumnado (siempre que la clase esté bien integrada y sin problemas especiales). Cuando haya algún sujeto con dificultades de integración, permite el conocimiento estructural del grupo, lo que servirá para incorporarlo con los compañeros que más le puedan motivar y con los que él se encuentre más compenetrado; tanto para el estudio como para el juego o cualquier otro tipo de actividad, resultando más fácil romper el aislamiento.

b. Profesional: Teniendo presente a Cataldi y otros [5] concordamos con su afirmación que, “debido al creciente desarrollo del software educativo durante los últimos años, gran parte del mismo ha sido realizado en forma desorganizada y poco documentada, y considerando el aumento exponencial que sufrirá en los próximos años, surge la necesidad de lograr una metodología disciplinada para su desarrollo, mediante los métodos,

procedimientos y herramientas, que provee la ingeniería de software para construir programas educativos de calidad, siguiendo las pautas de las teorías educativas y de la comunicación subyacentes”. Enmarcamos nuestra investigación dentro de los lineamientos actuales de la ingeniería del software [9], para obtener un software que sea de apoyo para funciones educativas, dado que no todos los programas de computación son específicos para educación.

c. Social: En esta tesis, se incorpora el ARS y el potencial que brinda para modelar los actores sociales, así como la diversidad y profundidad de las relaciones existentes entre alumnos. También se valoran una serie de estrategias y diversos instrumentos de medición para determinar las relaciones entre los alumnos, dentro de una realidad socioeducativa. La integración entre lo tecnológico y lo social, nos lleva a compartir el enfoque de la tecnología como sistema; considerando el concepto de red, como un buen descriptor de las relaciones, los actores inmersos en relaciones sociales que afectan y modifican su comportamiento y demás el ARS como un conjunto de técnicas de análisis para el estudio formal de las relaciones y estructuras sociales. A partir de diferentes escenarios sociales, se pueden obtener diferentes resultados que pueden ser de mucha utilidad al momento de la toma de decisiones.

5. TRABAJOS FUTUROS

Durante el desarrollo de esta tesis y a partir del desarrollo del software educativo Agent SocialMetric se han identificado problemas abiertos que pueden derivar las siguientes líneas de trabajo:

1. A través de los diferentes escenarios creemos que este trabajo puede orientar futuras investigaciones en las que el interés se centre en la interrelación entre los factores contextuales, ya que la complejidad de los mismos deja abierta la puerta a variables que no han sido incluidas en nuestro estudio. Por lo anterior, desde un enfoque cualitativo a futuro se pretende entrevistar y observar a los estudiantes en situaciones educativas mediadas con las TICs, con el objetivo de identificar nuevas estrategias en el uso de la herramienta [13].

Aprovechando la motivación que tienen los

alumnos por las TICs proponemos extender el trabajo de investigación a través de un módulo llamado Student Play [12] incorporando tres agentes conversacionales (llamados Max, Rodolfo y Javier) (ver Figura 7.1), que interactúen con los alumnos, sean compañeros de juegos para el aprendizaje y que brinden respuestas verbales y no verbales adecuadas. Además se buscará que a posteriori trabajen y se integren colaborativamente con el agente ya desarrollado Albert. El objetivo prioritario de este módulo será ofrecer diferentes recursos, estrategias y herramientas prácticas adaptadas a diferentes rangos etarios, aplicables a sesiones de clase programadas. Estará destinado a favorecer la promoción del aprendizaje de valores sociales y personales, a través de la comunicación e interacción [1] brindada por los agentes (ver Figura 6 el maquetado de algunas pantallas posibles de Student Play).



Figura 6. Prototipo Student Play con los agentes conversacionales Max Javier o Rodolfo

2. Se puede dar lugar a más investigaciones y a la integración de otras técnicas de IA. A lo largo del trabajo vimos que dos tipos de relaciones que afectan de sobremanera el clima en un aula de clases, son las relaciones de agresión y liderazgo entre pares de alumnos de un curso. Hoy sabemos que suelen estar presente en todas las escuelas y niveles y afecta no solo la calidad de las relaciones interpersonales y la convivencia escolar, sino también al propio desarrollo de los escolares directamente implicados. Es así como nace el bullying, éste es comúnmente definido como un maltrato entre pares donde se produce una persecución tanto física, como psicológica de parte de un estudiante contra otro, al que elige como víctima de ataques constantes. Esta acción negativa e intencionada, posiciona a las víctimas en situaciones de las que difícilmente pueden salir por sus propios medios. Es en este contexto donde este sistema puede tomar protagonismo,

en disminuir las oportunidades en las que se puede manifestar el acoso; a través de considerar la distribución espacial que tienen los alumnos dentro de la sala de clases, como un factor a considerar.

Dado que nuestra herramienta es integradora de varias técnicas, lo que buscaremos es implementar algoritmos genéticos para la generación de la distribución óptima de estudiantes, considerando diversas configuraciones posibles de las aulas de clase (mapeadas en las matrices que están modeladas en el sistema) en caso de identificarse una problemática de bullying a través de Agent SocialMetric. Considerando por ejemplo las aulas, que no poseen pasillos, en donde los alumnos están sentados uno inmediatamente después de otro, en todos los sentidos (la cual puede representar a la configuración que se utiliza comúnmente en las aulas universitarias que poseen sillas individuales), aulas que tengan un pasillo cada dos puestos y aquellas aulas que poseen un pasillo por cada puesto, es decir, se intercalan los pasillos con los puestos. El algoritmos genéticos (AG), son técnicas para búsquedas y optimización, basados en los principios de la genética y la selección natural. Estos algoritmos, hacen evolucionar una población de individuos (cromosomas), sometiendo a acciones aleatorias, semejantes a las que actúan en la evolución biológica (operadores genéticos: mutaciones y recombinaciones genéticas), así como también, a una selección de acuerdo con algún criterio, en función del cual se decide cuáles son los individuos más adaptados, que sobreviven, y cuáles los menos aptos, que son descartados. Los alumnos de una sala de clases en particular, pueden ser representados por un identificador único, de tal forma, que dos alumnos no estén asignados a un mismo puesto. El identificador se puede corresponder a un número natural correlativo, que parte desde el primer alumno representado en la sociomatrix, hasta el último de éstos. De tal manera que un cromosoma quede compuesto por los alumnos de un curso y por los pasillos de la sala de clase, representados con un valor por ejemplo como "0". Para definir la función objetivo ver la ecuación 1 (la cual puede

permitir evaluar cada distribución generada) se podría considerar que la pregunta hecha a los estudiantes, representa una relación negativa entre los individuos, de tal forma, que se pretende tener un puntaje lo menos negativo posible. En este caso x corresponde a un individuo, influencia es la función de evaluación, corresponde a la percepción negativa que tiene el curso sobre un alumno en particular. La función vecindario, corresponde a la valoración que un puesto tiene en la sala de clases. Y penalización corresponde a la penalización que se aplica, al posicionar un alumno agresor (alumno), cerca de sus víctimas.

$$\text{Fitness} = \sum_{i=1}^n (\text{Influencia}(x_i) * \text{Vecindario}(i) + \text{Penalización}(x_i)), \text{ con } N \leq |\text{alumnos}| \quad (1)$$

3. Se buscará explorar técnicas de simulación de diálogos y lograr una ampliación del gestor de diálogo integrándolo al contexto social del agente Albert (y adicionalmente a los agentes del módulo Studnet Play [12], destinado a los alumnos). En este caso se evaluarán diferentes modelos, como el modelo propuesto por Augello y otros [2] en el cual se sostiene que el diálogo es una actividad conjunta de procesos individuales y sociales. Para eso se explorará agentes sociales capaces de elegir los planes de diálogo, que sean más adecuados al contexto social (lo que en la literatura sociológica que se llama una "práctica social") donde se incorporen y analicen las habilidades comunicativas del agente. En esta línea se establecería una nueva la base de conocimientos, que incluya un conjunto de palabras que el agente reconozca modeladas mediante grafos gramaticales para establecer las relaciones entre las palabras claves; y un algoritmo de preprocesamiento que se pueda ajustar a cualquier frase ingresada y a un conjunto de palabras claves válidas.

Otra de las líneas será la posibilidad de expresar en los diálogos del agente Albert rasgos de su personalidad, emociones o intensidad (por ejemplo algunas de las emociones podrían ser alegría, sorpresa, resignación, etc.). Para incorporar las emociones a futuro se podrá establecer un modelo comportamental o emocional. Este módulo comportamental podrá contener un módulo de percepción e interpretación del contexto (el cual evalúe las acciones y actitudes del interlocutor), un módulo

de interacción social y emocional (el cual actuará según los rasgos de personalidad del personaje) y un módulo de gestión de diálogo (que pueda actualizar el estado del diálogo y seleccionar los actos del lenguaje que expresen los objetivos e intenciones del personaje).

Finalmente otra posible línea de experimentación podría partir de la base de explorar otros enfoques para la evaluación del diálogo hablado del agente Albert combinando diferentes apariencias (ej. humano en 3D, animado 2D, etc.) y distintas estrategias de comportamiento (ej. neutral, afectivo, etc.). De esta forma se podría analizar el compromiso entre las expectativas del agente y el funcionamiento del diálogo.

4. Aprovechando que la ingeniería de software educativo se viene convirtiendo en un área de estudio en expansión y dado que en nuestra investigación hemos implementado una metodología basada en Kanban es importante como línea futura considerar otras metodologías Ágiles para nuevos desarrollos y en la extensión de Agent SocialMetric (como en el módulo Student Play). Integrando y combinando componentes pedagógicos y didácticos a otras metodologías ágiles (como Extreme Programming XP, Crystal etc.) para el desarrollo de software educativo [8]. Así mismo se buscará contemplar en las distintas etapas y actividades aspectos de naturaleza interdisciplinaria desde una visión integradora.

5. En cuanto al ARS se propone como futura líneas incorporar redes dinámicas e implementar el “review extractor” de Agent SocialMetric . Por un lado, en cuanto a las redes dinámicas se prevé como futura línea incorporar un modelo que tenga en cuenta el paso del tiempo, en lugar de enfatizar la descripción de la totalidad de las relaciones, se buscará examinar el resultado de la decisión de los actores de establecer (mantener o romper) una relación teniendo en cuenta las relaciones preexistentes [6], su posible impacto en el comportamiento de los actores en el clima del aula y la posibilidad de explorar más métricas de redes. Por otro lado, a través del Review Extractor se buscará como línea futura estudiar el análisis léxico enfocado en Moodle y en Twitter. En cuanto a Moodle, el modulo Review Extractor

queda representado en la arquitectura de la Figura 8.

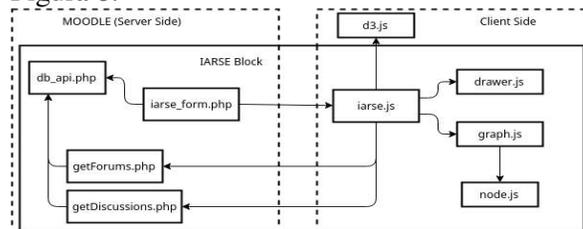


Figura 8. Arquitectura del módulo Review Extractor

El Review Extractor también podría trabajar con los tweets de Twitter, su relación con las tendencias de usuarios alumnos y cómo afecta al clima que se manifiesta en el aula en diferentes momentos. A través de la conexión con Twitter se buscará incluir el manejo de un clasificador de microtextos, para poder descubrir las características esenciales de los temas que se traten en los tweets seleccionados, ya sea de forma individual o masiva, ejecutándose en tiempo real o bien agregando el contenido a la base de datos que ofrece la aplicación Agent SocialMetric. Se podrían utilizar técnicas de tokenización para ayudar a procesar los tweets de forma más limpia y sin ruido, para la clasificación se podrían usar diccionarios que ayuden a identificar palabras (en formato XML) y el uso de algoritmos de clasificación como el de Naïve Bayes.

6. REFERENCIAS

- [1] Rubén Araujo, Caterina Clemenza, and Ramiro Fuenmayor. Las tecnologías de la información y de los medios de comunicación social como elemento clave en la educación en valores. *Encuentro Educativo*, 15(2), 2013.
- [2] Manueland Weideveld Lucasand Dignum Frank Augello, Agneseand Gentile. *A Model of a Social Chatbot*, pages 637–647. Springer International Publishing, Cham, 2016.
- [3] V. R. Basili. The Experimental Paradigm in Software Engineering. In *Proceedings of the International Workshop on Experimental Software Engineering Issues: Critical Assessment and Future Directions*, pages 3–12, London, UK, UK, 1993. Springer-Verlag.
- [4] M.A. Casanova. *La sociometría en el aula*. Aula abierta. La Muralla, 1991.

- [5] Zulma Cataldi, F. Lage, R. Pessacq, and R. García Martínez. Ingeniería de software educativo. In *Proceedings del V Congreso Internacional de Ingeniería Informática*, pages 185–199, 1999.
- [6] Ainhoa de Federico de la Rúa. El análisis dinámico de redes sociales con SIENA: método, discusión y aplicación. *Empiria: Revista de metodología de ciencias sociales*, (10):151–184, 2005.
- [7] P. Desberg, J.A. Colbert, and K.D. Trimble. *The Case for Education: Contemporary Approaches for Using Case Methods*. Allyn and Bacon, 1996.
- [8] A. Duarte Orjuela and Mauricio Rojas. Las metodologías de desarrollo ágil como una oportunidad para la ingeniería del software educativo. *Avances en Sistemas e Informática*, 5(2):159–172, 2008.
- [9] A. Galvis. *Ingeniería de software educativo*. Colección Biblioclase. Ediciones Uniandes, 2000.
- [10] A. A. García, R. De las Heras del Dedo, and C.L. Gómez. *Métodos Ágiles y Scrum*. Manuales Imprescindibles. Anaya Multimedia, 2012.
- [11] Theodore J. Kopcha. Teachers' Perceptions of the Barriers to Technology Integration and Practices with Technology Under Situated Professional Development. *Comput. Educ.*, 59(4):1109–1121, dec 2012.
- [12] A. Kuz, M. Falco, F. Castellini, and R. Giandini. Student Play: un Módulo Educativo de Agent SocialMetric. In *4to Congreso Nacional de Ingeniería en Informática/Sistemas de Información (CONAISI)*, 2016.
- [13] Antonieta Kuz, Mariana Falco, Roxana Giandini, and Leopoldo Nahuel. Agent SocialMetric: herramienta de asistencia al docente para determinar el clima social y la estructura del aula. *Revista Iberoamericana de Informática Educativa Comunicaciones (IE Comunicaciones)*, (22):16–29, Julio-Diciembre 2015.
- [14] L. Madueño. Desarrollo de Software Educativo bajo Plataforma Web. In *Congreso Internacional EDUTECH 2003, Gestión de las Tecnologías de la Información y la Comunicación en los Diferentes Ámbitos Educativos*, Venezuela, Noviembre 2003.
- [15] J. Martin and C.L. McClure. *Structured techniques for computing*. Prentice-Hall, 1985.
- [16] Gurpreet Singh Matharu, Anju Mishra, Harmeet Singh, and Priyanka Upadhyay. Empirical Study of Agile Software Development Methodologies: A Comparative Analysis. *ACM SIGSOFT Software Engineering Notes*, 40(1):1–6, feb 2015.
- [17] M. A. Moreira and Carina González. Líneas de investigación sobre Tecnologías de la Información. In *XI Jornadas Universitarias de Tecnología Educativa*, Universidad de Valladolid, Septiembre 2003.
- [18] M.A. Moreira. *Los medios y las tecnologías en la educación*. Colección psicología. Ediciones Pirámide, 2004.
- [19] Sung Hee Park and Peggy A Ertmer. Examining barriers in technology-enhanced problem-based learning: Using a performance support systems approach. *British Journal of Educational Technology*, 39(4):631–643, 2008.
- [20] (PASEM) Programa de Apoyo al Sector Educativo del Mercosur. *Incorporación con sentido pedagógico de TIC en la formación docente de los países del Mercosur*. Teseo, 2014.
- [21] Kevin Schoepp. Barriers to technology integration in a technology-rich environment. *Learning and teaching in higher education: Gulf perspectives*, 2(1):1–24, 2005.
- [22] M. Vidal Ledo, F. Gómez Martínez, and A.M. Ruiz Piedra. Software educativos. *Educación Médica Superior*, 24:97 – 110, 03 2010.
- [23] Michael Wooldridge. Editorial. *Autonomous Agents and Multi-Agent Systems*, 2(1):5, 1999.
- [24] Michael Wooldridge. Intelligent Agents: The Key Concepts. In *Multi-Agent-Systems and Applications*, pages 3–43, 2001.
- [25] M.V. Zelkowitz and D.R. Wallace. Experimental models for validating technology. *Computer*, 31(5):23–31, May 1998.

RESUMEN TESIS DOCTORAL

“ESTUDIO DE LA INFLUENCIA DE UN ENTORNO DE SIMULACIÓN EN LA ENSEÑANZA DE REDES DE COMPUTADORAS EN EL NIVEL UNIVERSITARIO”

Daniel Arias Figueroa

Director: Francisco Javier Díaz (UNLP)
Codirector: María Cecilia Gramajo (UNSa)

Área temática: Tecnología Aplicada en Educación
Fecha de defensa: 23 de noviembre de 2017
Jurado: Héctor Nelson Acosta, Eduardo Omar Sosa, Martín Antonio Navarro

Tesis presentada para obtener el grado de
Doctor en Ciencias Informáticas
Facultad de Informática – Universidad Nacional de La Plata

Resumen: Este artículo resume el trabajo de la tesis doctoral “Estudio de la influencia de un entorno de simulación en la enseñanza de redes de computadoras en el nivel universitario”, cuyo propósito fue determinar la influencia del uso de software de simulación en la enseñanza de conceptos y fundamentos sobre redes de computadoras, para un contexto específico. El estudio fue del tipo cuantitativo, con diseño experimental con grupo de control. Las pruebas paramétricas permitieron concluir que, existe diferencia estadísticamente significativa a favor de los estudiantes que emplearon una estrategia basada en simulación.

Palabras-claves: Enseñanza de redes de computadoras, Simulación, Laboratorio virtual, Tecnología Educativa, Protocolo TCP-IP.

1 INTRODUCCION

Los altos costos de equipos específicos necesarios para montar un laboratorio de red, sumados a los escasos recursos con los que cuenta el Departamento de Informática de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta (similares a los recursos con los que cuentan la mayoría de las universidades públicas en el interior del país), hacen considerar a las herramientas de simulación como una posible solución para que las prácticas sean mejor aprovechadas por los estudiantes, posibilitando además la utilización de estas herramientas fuera de los horarios de clases.

De acuerdo a lo dicho anteriormente, se podría resumir la problemática de la siguiente manera:

- El docente debe plantear los trabajos prácticos de laboratorio adecuándose a las características del equipamiento disponible, generalmente escaso (actualmente en el laboratorio del Departamento de Informática, existen 3 enrutadores marca Cisco, 5 enrutadores del tipo SOHO - Small Office Home Office y diversos conmutadores).

- La cantidad de estudiantes habitualmente es elevada (aproximadamente 15-20 estudiantes cada año).
- Los equipos de hardware (enrutadores, conmutadores, concentradores, cableado de red, conectores, etc.) son costosos, y su actualización y mantenimiento también significa costos elevados, por lo que usualmente se puede contar con a lo sumo uno o dos dispositivos por comisión o grupos de estudiantes. Lo que hace impracticable los laboratorios con equipo real.
- La curva de aprendizaje para la administración de los dispositivos es alta, lo mismo ocurre con la conectorización física para definir una determinada topología, ya que se disponen diferentes tipos de interfaces de red tal como Ethernet, FastEthernet y puertos Seriales. Esto impide realizar demasiados grupos que accedan al hardware de red.

1.1 Motivación

En muchas universidades la computadora, las redes, los sistemas de videoconferencias, la Internet y las soluciones de software se usan a diario. Toda esta infraestructura tecnológica no solo contribuye a facilitar las tareas y actividades como la documentación y obtención de información, si no que sirven como punta de lanza para la investigación sobre los cambios que se experimentan, tanto en el uso de herramientas en la educación como en los posibles métodos de enseñanza que se requieren para incorporar estos recursos.

Uno de los objetivos clave en la enseñanza de las redes de computadoras en la carrera de Licenciatura en Análisis de Sistemas es transmitir conceptos básicos y fundamentos a los estudiantes. Sin embargo, desde hace un tiempo se viene investigando como facilitar la relación entre la realidad y las teorías y modelos, es decir, entre lo concreto y lo abstracto.

Así, las computadoras personales (PC), con la variedad de software que se ha desarrollado, tienen en sí mismas un gran potencial para mejorar el proceso de enseñanza y aprendizaje, a fines que:

- El aprendizaje sea más interesante.
- El aprendizaje sea activo, no pasivo en las aulas.
- Los estudiantes estén más motivados.
- El aprendizaje sea al ritmo del estudiante en forma personalizada.
- La educación sea permanente.

1.2 Objetivos y preguntas de investigación

El objetivo general de esta tesis doctoral fue evaluar la influencia de la utilización de software de simulación en la enseñanza de contenidos de redes de computadoras en el ámbito de las asignaturas relacionadas con la temática en la carrera de Licenciatura en Análisis de Sistemas, cursos de extensión y postgrado organizados por el C.I.D.I.A. – Centro de Investigación y Desarrollo en Informática Aplicada que depende de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta – y en la carrera Ingeniería en Telecomunicaciones en la Universidad Católica de Salta. Para ello se llevó a cabo un estudio aplicado, longitudinal y experimental que, asumiendo la complejidad del fenómeno educativo, considera aspectos cuantitativos y cualitativos.

A partir de éste se plantearon los siguientes objetivos específicos:

- Establecer la relación entre la utilización de la simulación y la actitud de los estudiantes hacia el estudio de los protocolos IP;
- Determinar si existe diferencia significativa en el nivel de comprensión de los principios de los protocolos de red analizados, entre estudiantes que reciben instrucción mediada por herramientas de simulación y estudiantes que reciben instrucción tradicional.

Las preguntas que se plantearon en este sentido son:

- A. ¿Cómo influyen en la enseñanza y el aprendizaje los trabajos realizados con un software de simulación?, ¿Qué efectos causan los programas de simulación en lo que hace a la motivación del estudiante a aprender a resolver los problemas analizados en las clases teóricas?
- B. ¿Cuánto pueden contribuir los trabajos sobre el simulador a que se reafirmen los conceptos teóricos y la comprensión de los fundamentos y funcionamiento en general de las redes?
- C. ¿El trabajo con el simulador permite al estudiante sentir la seguridad de quien ha explorado un tema acabadamente, especialmente con la posibilidad de variar arbitrariamente los valores de los componentes y variables?
- D. Dado que el estudiante puede diseñar e inventar topologías y experimentar con ellas ¿hasta qué punto favorece la creatividad?
- E. ¿Hasta qué punto contribuyen a reemplazar un aprendizaje memorístico por aprendizaje significativo?
- F. ¿En qué momento conviene aplicar estos métodos, antes y/o después de los trabajos prácticos tradicionales? ¿Qué condiciones deben darse para la aplicación de estos métodos?
- G. ¿Es deseable estructurar otras actividades prácticas utilizando métodos de simulación?

Este estudio aportará evidencia empírica sobre la incidencia del uso de herramientas de simulación en la enseñanza de conceptos de redes de computadoras en el nivel universitario. Esta investigación beneficiará a los diferentes actores del proceso educativo: docentes, investigadores y autoridades educativas. Los resultados de esta investigación permitirán tomar importantes decisiones sobre la inclusión de herramientas TIC como medios para favorecer el aprendizaje de conceptos y fundamentos no solo de redes de computadoras.

2 ASPECTOS METODOLÓGICOS

2.1 Formulación de Hipótesis

De lo expuesto en el marco teórico, el resultado de la búsqueda realizada y la formación y experiencia adquiridas, surgieron las siguientes hipótesis causales de investigación, en relación a las correlaciones aprendizaje-métodos de simulación:

- **Hipótesis I:** A igualdad de recursos utilizados y experiencias, el estudiante aprende mejor, más significativamente, cuando complementa esas experiencias con métodos de simulación.

- **Hipótesis II:** El uso de software de simulación como complemento a los prácticos tradicionales o de laboratorio, mejora el aprendizaje de conceptos y la comprensión.
- **Hipótesis complementaria:** La realización de trabajos prácticos por simulación va a redundar en un aprendizaje significativo y creativo en los estudiantes. Esta hipótesis intenta responder a las preguntas C, D, y E del apartado: “objetivos y preguntas de investigación”.

Las hipótesis I y II intentan responder a las preguntas A, B, F y G del apartado: “objetivos y preguntas de investigación”.

2.2 Estrategia metodológica, diseño de la investigación y recolección de datos

La investigación se dividió en tres etapas: la primera etapa consistió en la aplicación de un cuestionario (inicialmente exploratorio), que se prolongó en todas las demás etapas. Por lo analizado en el marco teórico, cuando no se registran muchos antecedentes en relación a un tema, lo indicado es comenzar con un estudio de este tipo que permita preparar el terreno para la investigación posterior. Esta circunstancia definió el carácter y la profundidad con que se diseñó la investigación. A los efectos de precisar el lenguaje utilizado para caracterizar los estudios, se aclara que la terminología utilizada responde a la clasificación de Dankhe (1989), adoptada por Sampieri, (1998), quien los divide en: exploratorios, explicativos, descriptivos o correlacionales. En la segunda etapa se realizó un estudio explicativo-correlacional (evaluaciones de conceptos), y paralelamente un estudio descriptivo (encuesta para medir actitudes), que se prolongó en la tercera etapa.

Las actividades en las tres etapas se mencionan a continuación:

Primera Etapa	Segunda Etapa	Tercera Etapa
Cuestionario	Cuestionario	Cuestionario
Encuesta para medir actitudes	Encuesta para medir actitudes	Encuesta para medir actitudes
	Evaluación de conceptos	Evaluación de conceptos
		Encuesta a docentes

Tabla 2-1: Actividades en las distintas etapas de investigación

Para la segunda y tercera etapa, la investigación se diseñó de la siguiente forma: Se planificó realizar, por un lado, un estudio explicativo y correlacional a través de experimentos y por otro lado un estudio descriptivo a través de encuestas. En una primera instancia el estudio fue descriptivo, luego ambos estudios se realizaron en forma paralela, aunque temporalmente el estudio descriptivo fue realizado, cada año, al finalizar las experiencias con los trabajos prácticos propuestos.

2.2.1 Estudio explicativo-correlacional

Consistió en la realización de experiencias en las que se investigó la relación entre aprendizaje y los métodos de simulación y las actividades prácticas tradicionales. Las variables independientes, en este caso fueron:

- Trabajos prácticos tradicionales - TPT: en el aula.

- Actividades prácticas con simulación - TPS: con software de simulación.

Y como variables dependientes de ellas, a evaluar, se consideraron varios conceptos o protocolos como son:

- Protocolo ARP,
- Protocolo DNS,
- Direccionamiento IP,
- Ruteo IP: Estático y Protocolo RIP,
- Algoritmo CSMA/CD.

Estos y otros conceptos, el estudiante los profundiza con las actividades, y son los mismos, tanto en los TPT, como en los TPS.

Cabe aclarar también que el estudio y análisis de un protocolo o algoritmo de red no implica un solo concepto que puede ser memorizado, sino por el contrario un conjunto de reglas que necesariamente se tienen que dar para el funcionamiento del mismo.

En primer lugar, como parte de la experiencia, se impartieron los conceptos teóricos con la ayuda de presentaciones.

En la primera actividad práctica se utilizaron los conceptos relacionados al protocolo en estudio.

En la actividad práctica con software de simulación se trabajó con los mismos conceptos del trabajo práctico tradicional, solo que el estudiante necesitó familiarizarse antes en el uso del software específico para el tema en cuestión.

Dado que la cantidad de estudiantes en cada cohorte en las asignaturas en cuestión es pequeña, se dividió en dos grupos, un grupo experimental que realizó la práctica con software de simulación, y un grupo de control que realizó la práctica tradicional en el aula. La cantidad de estudiantes en cada grupo fue así función de la cantidad de estudiantes que constituyeron cada cohorte.

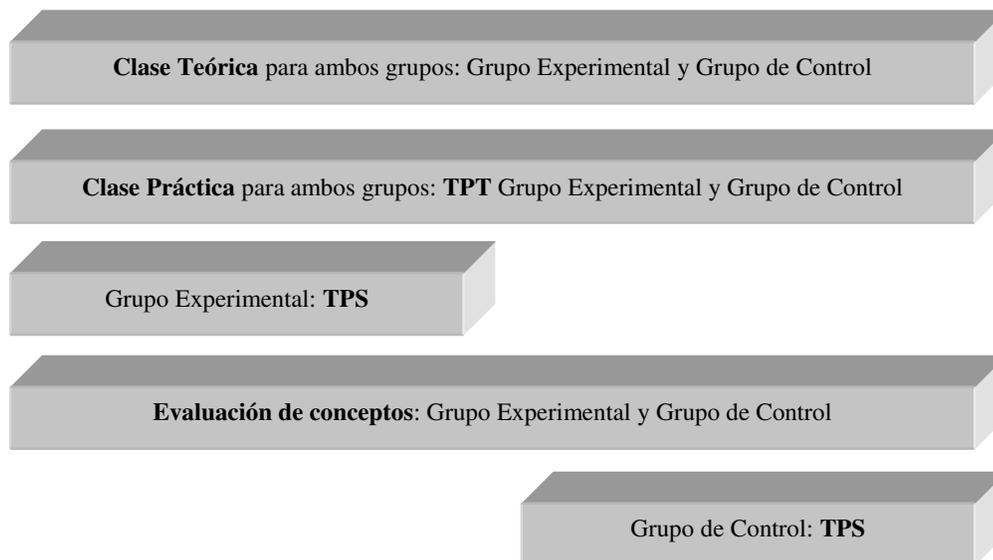


Figura 2-1: Distribución de las actividades, prácticas tradicionales y prácticas con simulación.

Este diseño cumple con todos los requisitos que debe reunir un experimento, a saber. Se considera por tanto que se han evitado todos los efectos que pueden quitarle validez interna a la experiencia.

En relación a la recolección de información a través de las evaluaciones, y por la forma en que fueron planeadas, se las considera confiables ya que el hecho de tener que aplicar el concepto a evaluar conlleva a responder siempre lo mismo en función de la estructura conceptual que tenga el estudiante en ese momento. Por la misma razón también son válidas, ya que no hay posibilidad de estar evaluando algo distinto a lo que se quiere evaluar. Las preguntas en todas las evaluaciones cubrieron completamente el contenido (los conceptos en cuestión a evaluar, protocolos y algoritmos). Por ser conceptos no tan definidos, resulta más complicado determinar en cada caso lo acertado, o no, de las respuestas (validez de criterio).

El momento de aplicación de las pruebas fue adecuado así como también la duración de las mismas. Se realizaron al comienzo de la clase ya que implicaba más concentración y actividad por parte del estudiante.

Las respuestas de las evaluaciones realizadas se puntuaron con una escala de 0 a 100 puntos. Las evaluaciones con repuestas correctas superiores a 50 puntos, se consideraron como una evaluación aprobada. En el caso de las evaluaciones de elección-múltiple que incluyeron una pregunta del tipo abierta, permitió conocer con mayor profundidad la claridad del concepto por parte del estudiante, y de esa manera poder detectar repuestas más elaboradas que otras.

Cabe aclarar que la evaluación de un protocolo o algoritmo implica el conocimiento de varios conceptos, no solo de uno.

La fiabilidad y validez de estos instrumentos fueron validadas a través de pruebas piloto donde se calcularon coeficientes de confiabilidad. Para el análisis de los datos se aplicó estadística descriptiva y pruebas paramétricas mediante el uso del Paquete Estadístico para las Ciencias Sociales (Statistical Package for the Social Sciences o SPSS® de IBM).

Para la prueba de conceptos, el objetivo principal fue contrastar estadísticamente si ambos grupos de estudio presentaron una diferencia significativa en cuanto a alguna medida de tendencia central o de variabilidad, a los fines de poder tomar una decisión confiable sobre el beneficio o no de la aplicación de la nueva técnica de enseñanza. Es decir, se pretendía determinar si el uso de la herramienta de simulación como complemento a las prácticas tradicionales, favorece el aprendizaje. Dado que los grupos de trabajo fueron relativamente pequeños, esto es, con un tamaño muestral chico ($N \leq 10$), para realizar las pruebas de contraste antes mencionadas en los casos que correspondiere, se utilizaron aquellas asociadas a la distribución normal o de Gauss. En los casos en que ello no fue posible, se recurrió a las técnicas contempladas en la estadística no paramétrica (ya que prescinden de la distribución de los datos).

La idea básica en el análisis estadístico de los datos fue utilizar teoría y el software mencionado para poder completar los resultados estadísticos descriptivos ya realizados, completándolos con estadísticas que permitan decidir por ejemplo: si existe diferencia significativa entre ambas metodologías de enseñanza.

Para cada experiencia se realizaron los test con pruebas específicas como Kolmogorov-Smirnov, coeficientes de correlación no paramétricos y gráficos específicos que permitieron ver características no siempre mostradas por los gráficos descriptivos. También se realizó la prueba de los rangos con signo de Wilcoxon que permitió contrastar la hipótesis nula de que ambos grupos puedan mantener algún tipo de relación, es decir, no fueren independientes los tratamientos de los métodos con y sin simulación.

2.2.2 Estudio descriptivo

El estudio descriptivo se realizó en la primera parte de la investigación y se prolongó en la segunda y tercera etapa. Para el test actitudinal se optó por una escala de valoración de Likert de 5 puntos para cada una de las variables o preguntas lo que permite medir actitudes y conocer el grado de conformidad del encuestado con las afirmaciones propuestas.

Para garantizar la fiabilidad del instrumento se utilizó el coeficiente Alfa de Cronbach, la cual asume que los ítems miden un mismo constructo y que están altamente correlacionados (Welch & Comer, 1988). Cuanto más cerca se encuentre el valor del alfa a 1, mayor es la consistencia interna de los ítems analizados. Como criterio general, George y Mallery (2003, p. 231) sugiere un valor de 7 como aceptable.

Los tests actitudinales indagaron sobre los siguientes puntos:

- Actitud de los estudiantes para realizar simulaciones con otra temática de la misma asignatura,
- Actitud de los estudiantes para realizar simulaciones en otras asignaturas del plan de estudio,
- Limitaciones de los simuladores,
- Forma en que influyen los cambios de determinados parámetros,
- Facilidad para experimentar con otras topologías,
- Aspectos específicos del simulador (facilidad de uso, ayuda contextual, facilidad de instalación, etc.),
- Claridad de las simulaciones,
- Apoyo brindado por la herramienta a la educación no presencial,
- Facilidad para realizar el seguimiento de los eventos,
- Implementación de los protocolos en el simulador,
- Tiempo asignado a los prácticos propuestos.

3 LA MUESTRA

Se realizaron 16 experiencias para el estudio descriptivo actitudinal. La Tabla 3-1, presenta para cada experiencia realizada la asignatura, la temática, el período y cantidad de estudiantes involucrados.

Experiencia en Asignatura	Temática	Período	Cantidad de Estudiantes
Conectividad y Teleinformática UNSa	Redes IP	2012	5
Conectividad y Teleinformática – UNSa	Redes IP, ARP, Fragmentación	2012	13
Redes de Computadoras I – UNSa	Protocolos Capa de Aplicación	2013	5
Curso DNS – UNSa	Protocolo DNS	2013	10
Redes de Computadoras I – UNSa	Fundamentos	2014	15
Redes de Computadoras I – UNSa	Protocolos Capa de Aplicación	2014	12
Redes de Computadoras I – UNSa	Protocolos Capa de Transporte	2014	9
Redes de Computadoras I – UNSa	Protocolos Capa de Red	2014	9
Curso Jornadas de Ingeniería	Ruteo con GNS3	2015	6
Curso de Extensión	Virtualización con GNS3	2015	6
Redes de Computadoras I - UNSa	Capas de Aplicación,	2015	8

Trasporte y Red			
Redes de Computadoras I - UNSa	Direccionamiento IP – Lab. real	2016	10
Redes de Computadoras I - UNSa	Ruteo IP – Lab. real	2016	9
Curso Mikrotik	Ruteo IP – Lab. real	2016	10
Encuesta UCASAL	Redes IP	2016	10
Encuesta docentes de redes	Redes IP	2016	13
Total e encuestados			140

Tabla 3-1: Muestra del estudio descriptivo actitudinal

Se trabajó con tamaños muestrales pequeños. Se diseñaron diferentes experiencias en diferentes asignaturas y diferente temática, muchas de las cuales se replicaron en años siguientes. La población de estudio estuvo conformada por la totalidad de los estudiantes cursantes en cada cohorte en cada una de las asignaturas bajo estudio.

En todos los casos se intentó que el grupo experimental y el grupo de control tuvieran la misma cantidad de estudiantes, lo que no siempre fue posible.

La Tabla 3-2, presenta las 12 experiencias que conforman la muestra para el estudio explicativo-correlacional de evaluación de conceptos, detallando la asignatura, la temática, el período y la cantidad de estudiantes que conformaron el grupo experimental y el grupo de control en cada caso.

Experiencia en Asignatura	Temática	Período	Grupo Experimental	Grupo de Control	Cantidad de Estudiantes
Conectividad y Teleinformática UNSa	Redes IP	2013	13	43	56
Redes de Computadoras II - UNSa	Redes LAN CSMA/CD	2014	4	5	9
Conectividad y Teleinformática – UNSa	Redes LAN CSMA/CD	2014	4	5	9
Redes de Computadoras II – UNSa	Protocolo ARP	2014	5	5	10
Redes I – UCASAL	Redes LAN CSMA/CD	2015	3	3	6
Redes de Computadoras II - UNSa	Redes LAN CSMA/CD	2015	5	7	12
Redes de Computadoras I - UNSa	Protocolo DNS	2015	5	16	21
Redes de Computadoras I - UNSa	Direccionamiento IP	2015	4	12	16
Redes de Computadoras I - UNSa	Ruteo IP	2015	5	11	16
Redes de Computadoras I - UNSa	Direccionamiento IP	2016	6	6	12
Redes de Computadoras I - UNSa	Ruteo IP	2016	5	8	13
Redes de Computadoras II - UNSa	Redes LAN CSMA/CD	2016	5	5	10
Total de estudiantes			64	126	190

Tabla 3-2: Muestra del estudio explicativo-correlacional

La muestra total para los dos tipos de experimentos asciende a un total de 330 estudiantes que participaron de la investigación.

Cabe destacar que los grupos bajo estudio pueden considerarse homogéneos o en igualdad de condiciones al inicio de cada experiencia, ya que se analizaron los planes de estudio correspondientes y el contenido de los programas de las respectivas asignaturas correlativas previas, y los mismos no contienen los conceptos evaluados en esta experiencia. Además no se incluyeron los estudiantes recursantes en ninguno de los grupos ni tratamientos, como se mencionó anteriormente.

4 RESULTADOS Y CONCLUSIONES

Los resultados alcanzados en esta investigación pretenden ser un aporte en la toma de decisiones de futuras propuestas curriculares y didácticas que tengan en cuenta los métodos de simulación, aplicados a contenidos concretos en el área de las redes de computadoras.

4.1 Hallazgos

Los resultados obtenidos en las Experiencias N° 1, 2, 3, 4, 5, 6, 10, 14, 16, 19 y 20, donde se trataron temas correspondientes a las Capa de Aplicación, Capa de Transporte y Capa de Red del Modelo TCP/IP, evidencian que los estudiantes consideran por una parte que las herramientas utilizadas (Kiva-NS, Packet Tracer y GNS3) cuentan con toda la potencialidad para realizar simulaciones claras de los protocolos estudiados y por otro lado consideran apropiados los laboratorios propuestos. De manera que, en relación con el primer objetivo del estudio, se establece que, para el nivel universitario, la utilización del software de simulación tiene una incidencia significativamente alta en el desarrollo de una buena actitud hacia el estudio de las asignaturas Redes de Computadoras I y Redes de Computadoras II del plan de estudios de la Licenciatura en Análisis de Sistemas.

En relación al segundo objetivo del estudio, y en función de los resultados obtenidos en las Experiencias N° 2, 6, 7, 8, 9, 11, 12, 13, 15, 17, 18 y 21, donde se realizaron evaluaciones de los conceptos Algoritmo CSMA/CD, Protocolo ARP, Protocolo DNS, Direccionamiento IP y Ruteo IP, se evidencia que el uso de software de simulación influye positivamente en el mejoramiento del nivel de comprensión de los conceptos. Se concluye en este aspecto que existe diferencia estadísticamente significativa a favor de los estudiantes que emplearon una estrategia basada en simulación.

Las Experiencias N° 22 y 23 permitieron contrastar la práctica con simulación y la práctica con equipo real, donde se evidencia que los estudiantes consideran que la simulación simplifica el proceso de configuración de topologías y dispositivos y el seguimiento de los eventos. Asimismo, consideran que la práctica con simulación puede reemplazar a la práctica con equipo real, cuando se trata de aprender fundamentos y conceptos sobre redes de computadoras.

Se considera que:

- las experiencias realizadas tienen “validez externa”, en cuanto a que pueden extrapolarse los resultados a otras asignaturas, de la misma carrera, que compartan contenidos y metodologías,
- los métodos de simulación tienen una influencia muy poderosa sobre la enseñanza, y un gran potencial a futuro.

4.2 Recomendaciones

El tema merece seguir siendo investigado por la importancia detectada en cuanto a su influencia en el aprendizaje. Se recomiendan estudios cualitativos que permitan conocer con mayor profundidad el nivel de conceptualización alcanzado por los estudiantes con el uso del software de simulación y ampliar la visión sobre el tema. También se sugiere la realización de otros experimentos que permitan contrastar el aprendizaje con simulación y el aprendizaje con equipo real considerando la enseñanza de conceptos y la formación de técnicos.

BIBLIOGRAFIA

- SAKAR, N. I. (2006) Teaching TCP/IP Networking Using Practical Laboratory Exercises, *International Journal of Information and Communication Technology Education*, Vol. 2, No. 4, pp. 39-50.
- GOLDSTEIN, G., M LEISTEN, S, STARK, K., & TICKLE, A. (2005) Using a Network Simulation Tool to Engage Students in Active Learning Enhances Their Understanding of Complex Data Communications Concepts, *Proceedings of the 7th Australasian conference on Computing Education*, pp. 223-228.
- JAVIDI, G. & SHEYBANI, E. (2008) Content-Based Computer Simulation of a Networking Course: An Assessment, *Journal of Computers*, Vol. 3, No. 3, pp. 64-72.
- DIXON, M. W., MCGILL, T. J. & KARISOON, J. M (1997) Using a Network Simulation Package to Teach the Client-server Model. *Proceedings of the 2nd Conference on Integrating Technology into Computer Science Education*, pp. 71-73.
- CAMERON, B. (2003): Effectiveness of simulation in a hybrid online networking course. *Quarterly Review of Distance Education*, 4(1), 51.
- GATTO, D. (1993): The use of interactive computer simulations in training. *Australian Journal of Educational Technology*, 9(2), 144-156.
- YAVERBAUM, G., & NADARAJAN, U. (1996): Learning basic concepts of telecommunications: an experiment in multimedia and learning. *Computers & Education*, 26(4), 215-224.
- ZHU, S. Y. (2011). Teaching Computer Networks through Network Simulation Programs. Faculty of Business, Computing and Law – School of Computing. University of Derby. *Learning Teaching & Assessment Conference*.
- KUROSE, J.F. & ROSS, K.W. (2015). Computer Networking: A Top-Down Approach. 6th Edition. *Pearson Education*. ISBN: 9780132856201.
- AVILA BLAS, Orlando José (2003). Probabilidad y estadística inferencial: teoría y aplicaciones. ISBN: 978-987-9381-23-6. *Editorial: Universidad Nacional de Salta*.
- GEORGE, D., & MALLERY, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn & Bacon.
- GLIEM, J & GLIEM, R. (2003). Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales. *Conference in Adult, Continuing, and Community Education. Midwest Research to Practice*.
- SIJTSMA, K. (2009). On the use, the misuse and the very limited of the Cronbach's Alpha. *Psychometrika*, 74(1) 107-120. DOI: 10.1007/S11336-008-9101-0
- CÁMARA, ALZUGARAY (2011). Trabajos Prácticos, Métodos de Simulación y aprendizaje significativo. *La Tecnología Educativa al servicio de la Educación Tecnológica*.
- SPAGNI, BEATRIZ (2007). Técnicas estadísticas aplicadas en la investigación con empleo del software específico S.P.S.S.". Material del curso dictado en UTN, Regional Santa Fe.
- GAMO, J. et al. (2015). Validación de requisitos funcionales de un Laboratorio Virtual Remoto como apoyo al blended learning. *Revista de Educación a Distancia*. 45(1).

Optimización de rendimiento, justicia y consumo energético en sistemas multicore asimétricos mediante planificación

Doctorado en Ciencias Informáticas
Facultad de Informática - Universidad Nacional de La Plata

Autor: **Adrian Pousa**

III-LIDI, Facultad de Informática, Universidad Nacional de La Plata
apousa@lidi.info.unlp.edu.ar

Fecha de exposición: 11 de Octubre de 2017

Directores:

Ing. Armando De Giusti

III-LIDI, Facultad de Informática, Universidad Nacional de La Plata
degiusti@lidi.info.unlp.edu.ar

Dr. Juan Carlos Sáez Alcaide

Dacya, Universidad Complutense de Madrid, España
jcsaezal@ucm.es

Resumen

Los procesadores multicore asimétricos (AMPs) fueron propuestos como una alternativa de bajo consumo energético a los procesadores multicore convencionales (CMPs) formados por cores idénticos. Los AMPs integran cores rápidos y complejos de alto rendimiento, y cores más simples de bajo consumo. Los AMPs plantean un gran desafío: distribuir eficientemente los ciclos de los distintos tipos de core entre las aplicaciones. Para evitar modificar el código de las aplicaciones, la mayoría de las propuestas se realizan a nivel de sistema operativo incluyendo algoritmos de planificación conscientes de la asimetría. Los algoritmos de planificación para AMPs propuestos hasta el momento intentan optimizar el rendimiento global pero degradan aspectos como la justicia o la eficiencia energética. Asimismo, la mayoría fueron evaluados mediante simuladores o plataformas asimétricas emuladas. El principal objetivo de esta tesis doctoral es superar estas limitaciones diseñando estrategias de planificación conscientes de la justicia y la eficiencia energética alcanzando un rendimiento global aceptable. Implementamos nuestras estrategias en el kernel de un SO real y las evaluamos sobre hardware multicore asimétrico real.

Palabras clave: multicore asimétrico, AMP, planificación, sistema operativo, rendimiento global, justicia, eficiencia energética

1. Introducción

La mayoría de los CMPs (*Chip Multi-Processors*) son simétricos, es decir, están compuestos por cores idénticos. Podemos encontrar CMPs formados por cores complejos (Haswell de Intel o Power8 de IBM) y CMPs formados por cores más simples de consumo reducido (ARM Cortex A9 o Intel Xeon Phi). Los procesadores del primer grupo poseen características microarquitectónicas sofisticadas (ejecución fuera de orden y superescalar) y son adecuados para aplicaciones secuenciales que los usan eficientemente. Los procesadores del segundo grupo son más simples en su microarquitectura y son adecuados para aplicaciones con un elevado paralelismo a nivel de hilo (TLP).

Podemos encontrar CMPs en varias plataformas donde se ejecutan cargas de trabajo muy diversas [13], que presentan distintas demandas y exigen al sistema operativo que optimice objetivos como el rendimiento global, la justicia o la eficiencia energética. Los CMPs presentan una limitación: un modelo particular de CMP puede resultar ideal, en cuanto a rendimiento por *watt*, para un conjunto de aplicaciones pero no para todas. Los AMPs fueron propuestos como alternativa a los CMPs para superar esta limitación [12, 2]. Un AMP está compuesto por dos tipos de core: cores rápidos de alto rendimiento (*big o fast cores*) y cores lentos de bajo consumo (*small o slow cores*). Todos los cores de un AMP poseen el mismo repertorio de instrucciones (ISA - *Instruction-Set Architecture*).¹

En un AMP es posible utilizar técnicas de *especialización*, es decir garantizar que una aplicación se ejecute en el tipo de core que ofrece la mejor relación entre rendimiento y consumo energético. En general, en un AMP los cores simples son adecuados para la ejecución de aplicaciones paralelas escalables[8]. Por el contrario, los cores complejos son adecuados para la ejecución de aplicaciones secuenciales. En el caso de aplicaciones paralelas que tienen fases secuenciales, es posible utilizar los cores rápidos para acelerar estas fases [1].

1.1. Motivación

A pesar de sus beneficios, los AMPs plantean importantes desafíos [6, 3, 13]. Uno de ellos es distribuir eficientemente los ciclos de los cores rápidos y lentos entre las distintas aplicaciones. Esta responsabilidad recae en el planificador del sistema operativo, de esta forma no es necesario modificar el código de las aplicaciones.

Al inicio de esta tesis doctoral, la mayoría de los algoritmos de planificación propuestos para AMPs tenían como objetivo optimizar el rendimiento global [12, 2, 10]. Para ello, el planificador debe ejecutar en los cores rápidos aquellas aplicaciones que usan estos cores de forma eficiente. La optimización de otros aspectos como la justicia o la eficiencia energética no habían recibido suficiente atención por parte de los investigadores. En esta tesis doctoral demostramos que los planificadores que intentan optimizar sólo el rendimiento global degradan estos otros aspectos.

¹Utilizar el mismo ISA permite ejecutar el mismo binario en los distintos cores sin tener que recompilar el código para cada uno de ellos.

1.2. Objetivos y principales desafíos

El principal objetivo de esta tesis es diseñar estrategias de planificación a nivel de sistema operativo para optimizar el rendimiento global, la justicia y la eficiencia energética. Para lograr este objetivo fue necesario superar tres grandes desafíos:

- Desarrollamos un *framework* de planificación que facilita la implementación y evaluación de las estrategias de planificación en un entorno realista: un sistema operativo real sobre hardware multicore asimétrico real. Por la complejidad que esto representa muchos investigadores evaluaron sus estrategias de planificación utilizando simuladores [li10, 2, 10, 16]. En esta tesis implementamos dichas estrategias sobre un SO real, esto nos permitió identificar ciertas limitaciones de algunas propuestas que no se manifiestan en entornos emulados[4].
- Equipamos al planificador del SO con un mecanismo para estimar en tiempo de ejecución el beneficio relativo (*speedup*) que una aplicación obtiene al usar los distintos tipos de cores en un AMP. Una aplicación puede obtener diferente *speedup* al ejecutarse en cores rápidos con respecto a hacerlo en los cores lentos [2]. Asignar aplicaciones a cores en base a esta diversidad de *speedups* es clave para mejorar el rendimiento global, la justicia o la eficiencia energética en AMPs. En una aplicación secuencial, el *speedup* se conoce como *Speedup Factor SF* y es el ratio del número de instrucciones por segundo (IPS) que el único hilo de la aplicación experimenta en ambos tipos de core ($\frac{IPS_{fast}}{IPS_{slow}}$). En aplicaciones paralelas, el *speedup* se obtiene a partir del *SF* de hilos individuales, el número de hilos de la aplicación y el número de cores rápidos. Medir directamente el *SF* da lugar a imprecisiones y genera *overhead* [15]. Por esta razón, optamos por estimar el *SF*. Para esto, construimos modelos de estimación de *SF* basados en contadores hardware.
- Seleccionamos métricas adecuadas para cuantificar el rendimiento global, la justicia y la eficiencia energética en un AMP. Las métricas que existían al inicio de la tesis estaban definidas para sistemas simétricos y no eran aptas para AMPs. Por este motivo, fue necesario construir nuevas métricas y adaptar métricas definidas previamente para CMPs.

1.3. Contribuciones de la tesis

Las principales contribuciones de la tesis doctoral son las siguientes:

- Construimos un modelo analítico para hallar los planificadores teóricos que optimizan el rendimiento global, la justicia y la eficiencia energética, respectivamente. A partir de éste modelo analizamos la interrelación entre estos tres aspectos y mostramos que no es posible optimizarlos simultáneamente.
- Proponemos los algoritmos de planificación para AMPs: Prop-SP, ACFS, EEF-Driven y ACFS-E. Prop-SP es la primera propuesta de planificación orientada

a justicia en AMPs que tiene en cuenta la diversidad de *speedups* entre aplicaciones. Aunque Prop-SP supera a otras estrategias de planificación, no es capaz de *optimizar* la justicia. Para superar esta limitación creamos ACFS, que maximiza la justicia en AMPs y permite ajustar gradualmente el nivel relativo de justicia y rendimiento global. EEF-Driven es el algoritmo que maximiza la eficiencia energética en AMPs en base al *Energy-Efficiency Factor (EEF)*. El *EEF* de un hilo es una nueva métrica propuesta en esta tesis que se define como $\frac{SF}{EPI_{fast}}$, donde EPI_{fast} es la energía por instrucción consumida por el hilo en el core rápido. ACFS-E es una variante de ACFS que maximiza la justicia, el rendimiento global y la eficiencia energética en un único algoritmo de planificación. Este planificador permite ajustar el nivel de eficiencia energética o rendimiento global a costa de degradar la justicia de forma gradual.

- Proponemos una metodología para construir modelos precisos de estimación de *SF* y *EEF* basados en el uso de contadores hardware.
- A diferencia de otras propuestas nuestras estrategias ofrecen soporte para acelerar distintos tipos de aplicaciones paralelas.

1.4. Publicaciones derivadas

- J. C. Saez, M. Prieto-Matías, A. Pousa, and A. Fedorova, "Explotación de técnicas de especialización de cores para planificación eficiente en procesadores multicore asimétricos", *Jornadas de paralelismo*, 2011.
- A. Pousa, J. C. Saez, A. D. Giusti, and M. Prieto-Matías, "Evaluation of scheduling algorithms on an asymmetric multicore prototype system", *XX Congreso Argentino de Ciencias de la Computación*, 2014.
- J. C. Saez, A. Pousa, F. Castro, D. Chaver, and M. Prieto Matías, "Exploring the throughput-fairness trade-off on asymmetric multicore systems", *Proc. of Euro-Par 2014: Parallel Processing Workshops - Euro-Par 2014 International Workshops, Part II*, 2014.
- J. C. Saez, A. Pousa, F. Castro, D. Chaver, and M. Prieto-Matías, "ACFS: A completely fair scheduler for asymmetric single-ISA multicore systems", *Proceedings of ACM/SIGAPP Symposium On Applied Computing*, 2015.
- J. C. Saez, A. Pousa, R. Rodriguez-Rodriguez, F. Castro, and M. Prieto-Matías, "PMCTrack: Delivering performance monitoring counter support to the OS scheduler", *The Computer Journal*, 2016.
- A. Pousa, J. C. Saez, F. Castro, D. Chaver, and M. Prieto-Matías, "Towards completely fair scheduling on asymmetric single-ISA multicore processors", *Journal of Parallel and Distributed Computing*, 2017.
- J. C. Saez, A. Pousa, A. D. Giusti, and M. Prieto-Matías, "On the interplay between throughput, fairness and energy efficiency on asymmetric multicore processors", *The Computer Journal*, 2017.

2. Entorno experimental

El entorno experimental utilizado en esta tesis está compuesto por el hardware asimétrico y el framework de planificación para AMPs.

Al inicio de esta tesis no existían plataformas asimétricas. Por esta razón, emulábamos el hardware asimétrico reduciendo la frecuencia de un conjunto de cores en un CMP para obtener los "cores lentos". Durante el desarrollo de la tesis surgieron AMPs con los cuales pudimos experimentar (Intel QuickIA y ARM Juno). Estas plataformas integran cores con diferencias en la frecuencia y en su microarquitectura.

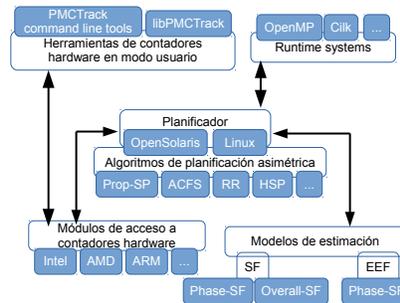


Figura 1: Framework de planificación.

La Figura 1 muestra la estructura del framework de planificación. Su componente central es el planificador del sistema operativo que interactúa con varios componentes que ayudan a la toma de decisiones de los algoritmos de planificación.

3. Métricas

La mayoría de las métricas existentes para cuantificar los distintos objetivos del planificador han sido definidas específicamente para CMPs y no resultan adecuadas para AMPs. Por lo tanto, definimos nuevas métricas y adaptamos métricas existentes.

Para cuantificar el rendimiento global en un AMP utilizamos el *Aggregate Speedup* (*ASP*), que se define como:

$$ASP = \sum_{i=1}^n \frac{CT_{Small,i}}{CT_{Sched,i}} - 1 \quad (1)$$

donde n es el número de aplicaciones en la carga de trabajo, $CT_{Small,i}$ es el tiempo de ejecución de la aplicación i cuando corre sola en el sistema y utiliza solamente los cores lentos, y $CT_{Sched,i}$ es el tiempo de ejecución de la aplicación i ejecutada bajo un planificador determinado junto al resto de aplicaciones de la carga de trabajo.

Para cuantificar la justicia adaptamos la métrica de *injusticia* o *unfairness*² propuesta por [14, 5], que se define como:

²Un algoritmo de planificación es justo si garantiza que las aplicaciones de una carga de trabajo experimentan la misma degradación del rendimiento (*slowdown*) al ejecutarse de forma simultánea con otras aplicaciones.

$$Unfairness = \frac{\max(Slowdown_1, Slowdown_2, \dots, Slowdown_n)}{\min(Slowdown_1, Slowdown_2, \dots, Slowdown_n)} \quad (2)$$

donde *Slowdown* para una aplicación *i* se define como $\frac{CT_{Sched,i}}{CT_{Fast,i}}$. $CT_{Fast,i}$ es el tiempo de ejecución de la aplicación *i* cuando corre sola en el sistema.

Para cuantificar la eficiencia energética en un AMP, utilizamos la métrica *Energy-Delay Product* (EDP)[9, 7], que se define como:

$$EDP = \frac{Energia_total_consumida \cdot CT}{Instrucciones_retiradas_totales} \quad (3)$$

donde *CT* es el tiempo de ejecución de la carga de trabajo (tiempo de ejecución de la aplicación más lenta).

4. Algoritmo de planificación Prop-SP

Al inicio de esta tesis, la mayoría de los algoritmos de planificación propuestos para AMPs perseguían maximizar el rendimiento global [12, 2, 10]. Sin embargo, maximizar el rendimiento degradaba la justicia. El objetivo de Prop-SP es ofrecer un buen equilibrio entre estos dos aspectos.

Prop-SP distribuye los ciclos de core rápido entre las aplicaciones utilizando una estrategia basada en créditos inspirada en el planificador Credit Scheduler de Xen sobre CMPs. Cada hilo tiene asociado un contador de *créditos* de core rápido. Aquellos hilos que poseen *créditos* pueden ejecutarse en este tipo de cores. Cuando un hilo se ejecuta en un core rápido, sus créditos se van consumiendo. Cada cierto tiempo, Prop-SP inicia un proceso que otorga créditos de core rápido a las aplicaciones activas teniendo en cuenta sus *speedups*. Cuando un hilo que está ejecutando en un core rápido agota todos sus créditos, el planificador intentará intercambiarlo con otro hilo que esté asignado a un core lento y posea créditos de cores rápidos.

Prop-SP mejora las propuestas de otros autores pero no optimiza la justicia.

5. Relación rendimiento, justicia y eficiencia energética

Realizamos un estudio teórico donde evaluamos la efectividad de diferentes algoritmos de planificación en cuanto al rendimiento global, la justicia y la eficiencia energética. A partir de este estudio analizamos la interrelación entre estos aspectos. Para esto, utilizamos varias cargas de trabajo (Cuadro 2) y las analizamos sobre un AMP compuesto por dos cores rápidos y dos cores lentos.

Cada carga de trabajo está formada por cuatro aplicaciones secuenciales de la suite de benchmarks SPEC CPU (Cuadro 1). Para construir las cargas de trabajo analizamos el comportamiento de varias aplicaciones cuando se ejecutan sobre la placa ARM Juno. Para obtener los valores de IPS y EPI utilizamos los contadores hardware y los registros de energía integrados en esta placa.

Analizamos los valores de las métricas *ASP*, *unfairness* y *EDP* al ejecutar las cargas de trabajo bajo cuatro estrategias de planificación para AMP:

Cuadro 1: Aplicaciones sintéticas

App.	Benchmark	IPS _{fast}	IPS _{slow}	SF	EPI _{fast}	EPI _{slow}
A1	art	0.60	0.24	2.47	1.59	1.86
A2	astar	0.58	0.31	1.86	1.40	1.10
A3	bzip2	1.49	0.73	2.02	0.61	0.45
A4	equake	0.80	0.26	3.07	1.31	1.45
A5	galgel	1.30	0.41	3.16	0.82	0.91
A6	gamess	2.01	0.69	2.91	0.51	0.49
A7	gobmk	1.09	0.61	1.79	0.75	0.54
A8	gzip	1.07	0.63	1.70	0.78	0.56
A9	h264ref	1.83	0.93	1.96	0.51	0.37
A10	hmmer	2.80	1.04	2.69	0.42	0.36
A11	mcf	0.18	0.09	2.02	3.98	4.44
A12	mgrid	1.28	0.59	2.17	0.85	0.65
A13	perlbench	1.42	0.71	2.01	0.60	0.47
A14	perlbmk	1.77	0.78	2.27	0.54	0.41
A15	povray	1.15	0.53	2.19	0.85	0.66
A16	soplex	0.52	0.20	2.53	1.68	1.86
A17	swim	0.25	0.11	2.24	3.11	3.34
A18	vortex	1.73	0.72	2.41	0.56	0.45
A19	wupwise	1.63	0.64	2.56	0.66	0.54

Cuadro 2: Cargas de trabajo

Carga de trabajo	Aplicaciones
W1	A5,A4,A6,A10
W2	A16,A17,A13,A9
W3	A16,A1,A18,A14
W4	A5,A4,A10,A15
W5	A5,A4,A6,A12
W6	A1,A12,A3,A13
W7	A1,A17,A7,A8
W8	A15,A11,A13,A2
W9	A4,A11,A3,A8
W10	A10,A19,A16,A9

- HSP: planificador que optimiza el rendimiento asignando a cores rápidos aquellas aplicaciones de la carga de trabajo con mayor *speedup*.
- *Opt-Unf*: planificador teórico que optimiza la justicia (asegura el óptimo *unfairness* - valor más bajo - para el máximo valor de *ASP* alcanzable).
- *Opt-EDP*: planificador teórico que optimiza la eficiencia energética (asegura el óptimo *EDP* -valor más bajo- para el máximo valor de *ASP* alcanzable).
- *EEF-Driven*: nuestro planificador que asigna a cores rápidos las aplicaciones que alcanzan el valor más alto de *EEF* (las restantes se asignan a cores lentos).

Las Figuras 2a y 2b muestran la relación entre rendimiento global (*ASP*) y eficiencia energética (*EDP*), y justicia (*unfairness*) y eficiencia energética, respectivamente. HSP alcanza los valores más altos de *ASP* a expensas de degradar la eficiencia energética. *Opt-EDP* alcanza el menor valor de *EDP* pero degrada significativamente el rendimiento global. Asimismo, tanto HSP como *Opt-EDP* son inherentemente injustos. Por otro lado, *Opt-Unf* alcanza los mejores valores de *unfairness* pero degrada tanto el rendimiento global como la eficiencia energética.

La principal conclusión que obtenemos de estos resultados es que la justicia, el rendimiento global y la eficiencia energética son objetivos contrapuestos, ya que cualquier intento de optimizar una de las métricas lleva a degradar las otras.

Opt-Unf y *Opt-EDP* no pueden ser implementados en el kernel de un SO, debido a que son algoritmos de orden exponencial. Observamos que *EEF-Driven* aproxima a *Opt-EDP*, ya que ambas realizan la misma distribución de ciclos de core rápido.

6. Algoritmo de planificación ACFS

ACFS intenta aproximar el planificador óptimo de justicia (*Opt-Unf*) mediante el seguimiento del progreso realizado por los distintos hilos de la carga de trabajo durante su ejecución. Para seguir el progreso de las distintas aplicaciones, el planificador

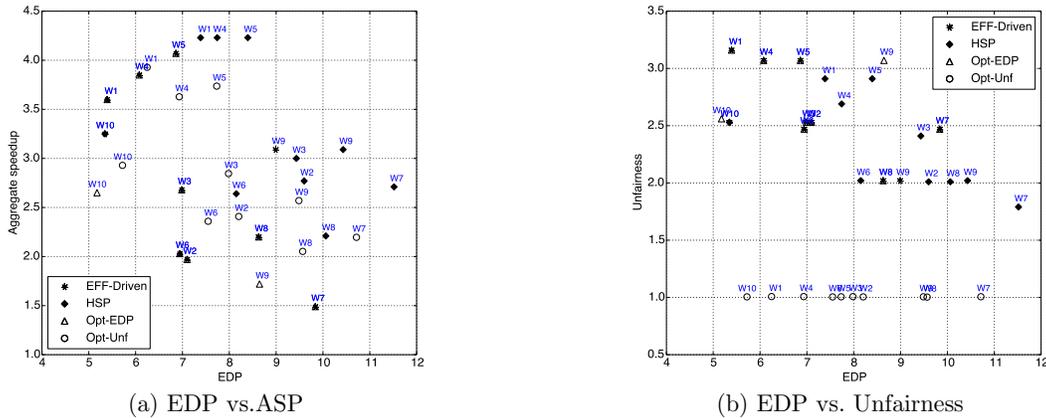


Figura 2: ASP, EDP y Unfairness para las cargas de trabajo del Cuadro 2 bajo distintos planificadores. Los mejores resultados en a) se encuentran próximos a la esquina superior izquierda y en b) se encuentran próximos a la esquina inferior izquierda.

asocia a cada hilo un contador ($amp_vruntime$). Cuando un hilo se ejecuta durante un *tick* de reloj en un tipo de core determinado, ACFS incrementa el contador del hilo en $\Delta amp_vruntime$, que se calcula como sigue:

$$\Delta amp_vruntime = \frac{100 \cdot W_{def}}{S_{core} \cdot W_t} \quad (4)$$

donde W_{def} es el peso de las aplicaciones con la prioridad por defecto, S_{core} es la degradación en rendimiento (*slowdown*) experimentada por la aplicación y W_t es el peso del hilo (prioridad especificada por el usuario).

Cuando un hilo se ejecuta en un core rápido, $S_{core} = 1$ (no hay degradación del rendimiento). Si el hilo se ejecuta en un core lento $S_{core} = Speedup$.

ACFS aproxima el *Speedup* en tiempo de ejecución teniendo en cuenta el *SF* del hilo y el número de hilos activos de la aplicación. A su vez, el *SF* del hilo se estima alimentando un modelo de estimación con valores de métricas de rendimiento obtenidos para el hilo mediante contadores hardware.

Para garantizar justicia, el planificador debe lograr que las aplicaciones realicen el mismo progreso. Para esto, ACFS garantiza que el valor de los contadores de progreso de los hilos sea lo más cercano posible, realizando migraciones de hilos entre los diferentes tipos de core cada cierto tiempo. ACFS efectúa intercambios de hilos cuando detecta que la diferencia entre los contadores de progreso ($amp_vruntime$) de estos supera un cierto umbral.

ACFS está equipado con un parámetro de configuración, *Unfairness Factor* o *UF*, que permite incrementar gradualmente el rendimiento global del sistema en escenarios donde los requisitos de justicia son menos estrictos.

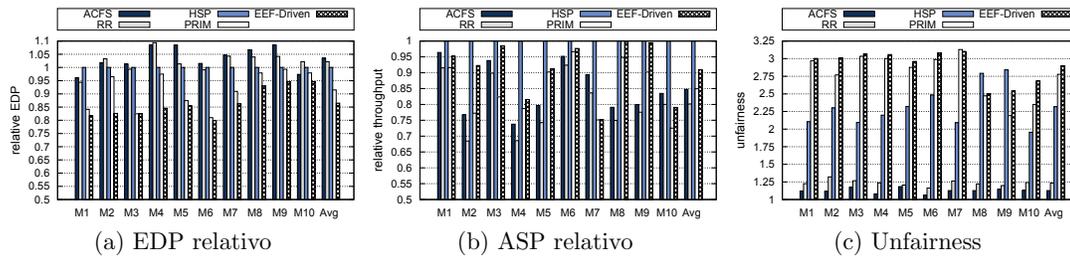


Figura 3: Resultados para las cargas de trabajo bajo la configuración 2F-4S

7. Análisis experimental

Comparamos EEF-Driven y ACFS con otros planificadores para AMP: HSP, PRIM [16] (orientado a mejorar la eficiencia energética) y RR (orientado a proporcionar justicia). Utilizamos la placa de desarrollo ARM Juno que integra 2 cores rápidos y 4 cores lentos (2F-4S). Evaluamos 10 cargas de trabajo (M1-M10), cada una compuesta por 6 aplicaciones de la suite de benchmarks SPEC CPU.

La Figura 3 muestra los valores de EDP , ASP y $Unfairness$ obtenidos. Los valores de EDP y ASP están normalizados con respecto a los resultados de HSP , y las cargas de trabajo están ordenadas por el EDP relativo dado por EEF-Driven.

Los resultados experimentales muestran tendencias similares al estudio teórico: optimizar una métrica puede llevar a degradar las otras. HSP alcanza los mejores valores de ASP pero sufre de degradación en EDP . EEF-Driven obtiene los mejores valores de EDP pero sufre de degradación en rendimiento. Ambos planificadores sufren de degradación en la justicia.

ACFS obtiene mejoras en justicia pero degrada el EDP y el ASP . ACFS supera a RR en rendimiento global y justicia. Esto se debe a que ACFS tiene en cuenta la *speedup* de las aplicaciones al tomar decisiones de planificación, a diferencia de RR.

EEF-Driven supera a PRIM en varios aspectos. En particular, PRIM hace intercambio de hilos de forma aleatoria, lo que provoca asignaciones de hilos a cores subóptimas y esto lo lleva a obtener valores de EDP peores a los esperados.

8. Algoritmo de planificación ACFS-E

Las estrategias de planificación evaluadas anteriormente ofrecen un compromiso fijo entre justicia, rendimiento global y eficiencia energética. Dado que estos aspectos no se pueden optimizar simultáneamente diseñamos ACFS-E. Este planificador, variante de ACFS, está equipado con dos parámetros de configuración: EDP_factor y $Unfairness_factor$ (UF). ACFS-E permite modificar la relación eficiencia energética-justicia mediante EDP_factor , o rendimiento global-justicia mediante UF . Cuando estos parámetros se establecen a sus valores por defecto, el algoritmo se comporta como la implementación base del algoritmo ACFS, que maximiza la justicia.

Analizamos cuatro cargas de trabajo bajo ACFS-E. Los resultados se muestran en las Figuras 4a y 4b. La Figura 4a muestra el impacto sobre la justicia y la efi-

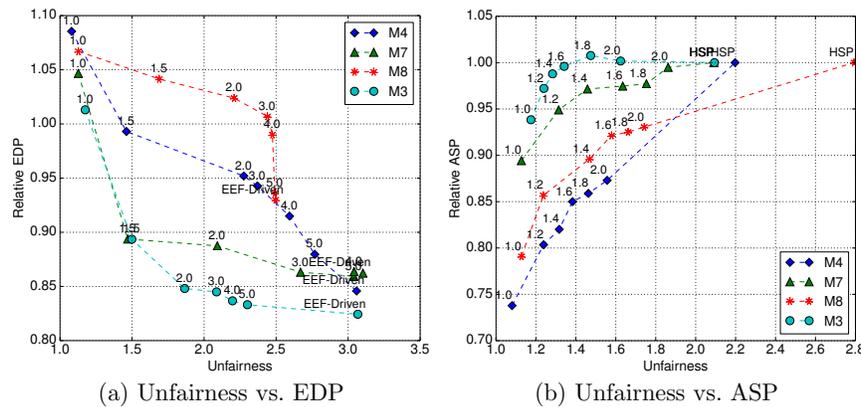


Figura 4: a) EDP Factor y b) Unfairness Factor.

ciencia energética al variar EDP_factor . En general, observamos que valores altos de EDP_factor tienden a reducir el EDP degradando la justicia. Así, ACFS-E se aproxima a EEF-Driven. La Figura 4b muestra el impacto sobre la justicia y el rendimiento global al variar UF . En general, observamos que valores altos de UF mejoran el rendimiento global degradando la justicia. Así, ACFS-E se aproxima a HSP.

Conclusiones

El principal objetivo de esta tesis doctoral ha sido diseñar estrategias de planificación para AMPs sobre un sistema operativo real, conscientes de la justicia y de la eficiencia energética, garantizando un rendimiento global aceptable. Con este objetivo, analizamos la relación entre estos tres aspectos y concluimos que son objetivos contrapuestos y no pueden ser optimizados simultáneamente en un AMP.

Este análisis fue clave para el diseño de nuestras estrategias de planificación: Prop-SP, ACFS, EEF-Driven y ACFS-E. El objetivo de Prop-SP es ofrecer un buen equilibrio entre justicia y rendimiento global. Sin embargo, este algoritmo no es capaz de optimizar la justicia en un AMP. Para superar esta limitación diseñamos ACFS, un algoritmo que aproxima el comportamiento del planificador teórico que optimiza la justicia en un AMP. Además, permite ajustar gradualmente el nivel relativo de justicia y rendimiento global. EEF-Driven aproxima el planificador teórico que optimiza la eficiencia energética en AMPs. ACFS-E es una variante de ACFS que constituye la primera estrategia de planificación para AMPs que puede configurarse para optimizar el rendimiento global, la justicia y la eficiencia energética individualmente, con un único algoritmo de planificación. Asimismo, ACFS-E permite al usuario ajustar el compromiso rendimiento-justicia o energía-justicia.

En el análisis experimental, mostramos que nuestros algoritmos superan estrategias propuestas por otros autores.

Referencias

- [1] Murali Annavaram, Ed Grochowski y John Shen. “Mitigating Amdahls Law through EPI Throttling”. En: Proc. of ISCA 05 (2005), págs. 298-309.
- [2] Michela Becchi y Patrick Crowley. “Dynamic Thread Assignment on Heterogeneous Multiprocessor Architectures”. En: Proc. of CF 06 (2006), págs. 29-40.
- [3] N. Chitlur y col. “QuickIA: Exploring heterogeneous architectures on real prototypes”. En: HPCA 12 (2012), págs. 1-8.
- [4] Kenzo Van Craeynest y col. “Fairness-aware scheduling on single-ISA heterogeneous multi-cores”. En: PACT 13 (2013), págs. 177-187.
- [5] Eiman Ebrahimi y col. “Fairness via source throttling: a configurable and high-performance fairness substrate for multi-core memory systems”. En: ASPLOS 10 (2010).
- [6] Matt Gillespie. “Preparing for The Second Stage of Multi-Core HW: Asymmetric (Heterogeneous) Cores”. En: *Intel White Paper* (2008).
- [7] R. Gonzalez y M. Horowitz. “Energy dissipation in general purpose microprocessors.” En: Solid-State Circuits, IEEE Journal (1996), págs. 1277-1284.
- [8] M. D. Hill y M. R. Marty. “Amdahls Law in the Multicore Era”. En: *IEEE Computer* (2008), págs. 33-38.
- [9] M. Horowitz, T. Indermaur y R. Gonzalez. “Low-power digital design.” En: Digest of Technical Papers., IEEE Symposium (1994), págs. 8-11.
- [10] David Koufaty, Dheeraj Reddy y Scott Hahn. “Bias Scheduling in Heterogeneous Multi-core Architectures”. En: Proc. of Eurosys 10 (2010).
- [11] Farkas Kumar y Jouppi. “Single-ISA Heterogeneous Multi-Core Architectures: the Potential for Processor Power Reduction”. En: Proc. of MICRO 36, 03 (2003).
- [12] Rakesh Kumar y col. “Single-ISA Heterogeneous Multi-Core Architectures for Multithreaded Workload Performance”. En: ISCA 04 (2004), págs. 64-.
- [13] Sparsh Mittal. “A Survey Of Techniques for Architecting and Managing Asymmetric Multicore Processors”. En: *ACM Computing Surveys* (2016).
- [14] Onur Mutlu y Thomas Moscibroda. “Stall-Time Fair Memory Access Scheduling for Chip Multiprocessors”. En: Proc. of MICRO '07 (2007). DOI: <http://dx.doi.org/10.1109/MICRO.2007.40>.
- [15] Daniel Shelepov y col. “HASS a Scheduler for Heterogeneous Multicore Systems”. En: *ACM SIGOPS OSR* 43.2 (2009).
- [16] Y. Zhang y col. “Cross-architecture prediction based scheduling for energy efficient execution on single-isa heterogeneous chip-multiprocessors.” En: *Microprocess. Microsyst.*,39 (2015), págs. 271-285.

Un enfoque inteligente para la selección de grupos de expertos mediante redes sociales

Dr. Eduardo Zamudio (*autor*)

Prof. Dra. Analía A. Amandi (*director*)

Prof. Dr. Luis S. Berdún (*co-director*)

Tesis doctoral defendida el 10 de marzo de 2017

Doctorado en Ciencias de la Computación, FCE, UNICEN

Resumen

La selección de grupos de expertos generalmente considera la evaluación de los criterios de selección que deben cumplir los candidatos y los grupos en sí mismos. Los criterios de selección representan un aspecto clave de los procesos de selección de expertos, ya que son estos criterios los que establecen qué candidato es un experto, o cómo se debe conformar un grupo de expertos.

En esta tesis se propone un enfoque inteligente e integral para el problema de la selección de grupos de expertos compuesto por dos partes. Primero, se propone un método para la identificación y evaluación de criterios de selección de candidatos elegibles a ocupar posiciones de expertos, utilizando descripciones de los candidatos como fuente de información. Segundo, se propone un método para la selección óptima de conformaciones de grupos de expertos, utilizando información relacional de los candidatos.

Esta tesis introduce importantes contribuciones en el área de selección de expertos, incluyendo la aplicación de Aprendizaje Automático (ML) en la identificación de evidencia de experiencia; la aplicación de técnicas de Procesamiento de Lenguaje Natural (NLP) para determinar la similitud de criterios de selección; una nueva métrica de Análisis de Redes Sociales (SNA) para determinar la independencia de grupos; la aplicación de una estrategia multicriterio para la evaluación de grupos en redes sociales; y la implementación de un algoritmo evolutivo para la selección óptima de grupos de expertos.

Las evaluaciones experimentales indican que la elegibilidad de un conjunto de candidatos puede

ser determinada a partir del nivel de correspondencia semántica entre las evaluaciones de los candidatos y los criterios de selección de referencia. Asimismo, los resultados indican que es posible recomendar grupos de expertos con mejor desempeño al compararlos con los comités actuales, a partir del uso de información relacional.

1. Introducción

Un grupo de expertos es un conjunto de personas con reconocido dominio de uno o varios temas. Los grupos de expertos se constituyen por varios motivos, como el tratamiento de asuntos específicos, definición de políticas, asignación de recursos, y asesoramiento, entre otros.

Por lo general, los grupos de expertos son seleccionados mediante algún proceso de selección de grupos, en el cual se evalúan tanto los requisitos que deben cumplir los candidatos, como así también los requisitos de los grupos en sí mismos. Estos requisitos se definen a partir de los criterios de selección que, cuando están disponibles, se pueden encontrar definidos en los lineamientos generales de los mismos procesos de selección.

La falta de acceso a los criterios de selección, o la subjetividad con la que éstos son definidos, limitan la posibilidad de realizar análisis objetivos de dichos procesos. Esto implica un desafío, principalmente en relación con las fuentes de información disponibles para realizar cualquier análisis sobre procesos de selección.

Por una parte, en varios procesos de selección de expertos, la única fuente de información disponible se constituye a partir de las evaluaciones de los candidatos. Estas evaluaciones son documentos de texto en los que los evaluadores argumentan la adecuación del candidato a los criterios de selección.

Por otra parte, la selección de grupos de expertos difícilmente se acompañe de un documento que describa la evaluación de los criterios de selección del grupo. En estos casos, la primera fuente de información disponible se constituye a partir del listado de miembros del grupo seleccionado.

En esta tesis se propone un enfoque integral para el problema de la selección de grupos de expertos, el cual consiste de dos aspectos. Por una parte, se propone determinar la elegibilidad de candidatos a ocupar posiciones de experto, mediante un método para la identificación y evaluación de los

criterios de selección utilizados. Por otra parte, se propone un método para la selección óptima de conformaciones de grupos de expertos, basada en las relaciones sociales de candidatos elegibles.

La hipótesis general del trabajo sostiene que es posible desarrollar un método objetivo de selección de grupos de expertos utilizando técnicas del área de la inteligencia artificial aplicada sobre un conjunto de candidatos para la conformación de dichos grupos.

Esta tesis introduce importantes contribuciones en el área de selección de expertos, incluyendo:

- Un enfoque de aprendizaje supervisado para la extracción de descripciones de candidatos, como una nueva fuente de evidencia de experiencia para procesos de selección de expertos.
- Una nueva aplicación de técnicas de Procesamiento de Lenguaje Natural (NLP) para determinar la similitud del contenido semántico de criterios de selección de expertos.
- Una nueva métrica de Análisis de Redes Sociales (SNA) para determinar la independencia de miembros de grupos [17, 18, 19].
- Un método de selección óptima de grupos a partir de la integración de métricas de selección de grupos basadas en redes sociales y la implementación de un algoritmo genético.

Este documento se organiza de la siguiente manera. En la sección 2 se presenta una visión general del método propuesto para la selección inteligente de grupos de expertos, separados en las etapas de (1) identificación de candidatos elegibles para la conformación de grupos de expertos, y (2) la representación del problema de selección de grupos mediante redes sociales, junto con el método propuesto para la optimización de conformaciones posibles de grupos de expertos. En la sección 3 se presentan las evaluaciones experimentales. En la sección 4 se presentan los antecedentes del trabajo. Finalmente, en la sección 5 se presentan las conclusiones.

2. Visión general de la propuesta

El enfoque inteligente e integral propuesto para el problema de la selección de expertos consiste de dos partes. Por una parte, con objeto de determinar la elegibilidad de candidatos a ocupar po-

siciones de experto, se propone un método para la identificación y evaluación de criterios de selección aplicados en la evaluación de dichos candidatos. Por otra parte, con objeto de seleccionar grupos de expertos, se propone un método para la selección óptima de conformaciones, basada en las relaciones sociales de sus miembros.

La figura 1 presenta el enfoque general de la propuesta con sus procesos principales. El proceso de elegibilidad de candidatos toma información de los candidatos a conformar los grupos de expertos, junto con información de los criterios de selección. La salida de este proceso es un subconjunto de candidatos elegibles, los cuales cumplen con los criterios del proceso de selección de expertos. Luego, el proceso de selección de grupos de expertos toma la información del subconjunto de candidatos elegibles, y de las relaciones sociales de éstos, con objeto de determinar las mejores conformaciones posibles.

2.1. Elegibilidad de candidatos mediante criterios de selección de expertos

Las evaluaciones de los candidatos, cuando son expresadas en lenguaje natural, suelen contener *descripciones*, las cuales se definen en este trabajo como las características destacadas de los candidatos desde el punto de vista de la persona responsable de una evaluación. Una descripción en un contexto de selección de expertos podría ser: “The candidate have demonstrated skills for conflict resolution”.

Considerando que los candidatos de un proceso de selección de expertos deben ser evaluados en base a un conjunto de criterios de selección, es posible asumir que los criterios de selección (explícitos o no) se encuentran implícitos en las descripciones de los candidatos contenidas en sus evaluaciones. En forma complementaria, puede ocurrir que los criterios de selección se encuentren explícitamente, por ejemplo, en guías, estatutos, lineamientos, etc. Se establece entonces una distinción entre los criterios de selección *explícitos* e *implícitos*.

Por lo tanto, resulta necesario un mecanismo que permita determinar en qué medida se asemejan dos conjuntos de criterios de selección, uno correspondiente a criterios de referencia (como los explícitos) y otro correspondiente a los criterios de correspondencia (como los implícitos).

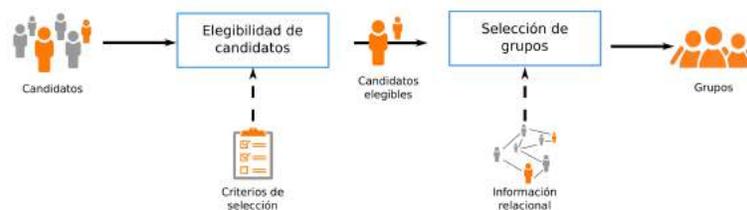


Figura 1: Enfoque general propuesto de selección de grupos de expertos.

Esta situación conduce a la siguiente hipótesis:

Los criterios de selección explícitos e implícitos (de un mismo proceso de selección) presentan un contenido semántico similar entre sí.

De esta manera, sería posible analizar la similitud semántica entre criterios (explícitos e implícitos), y así determinar en qué medida se aplican estos criterios en la evaluación de candidatos. Así, la *elegibilidad* de un candidato estaría dada cuando los criterios implícitos en sus evaluaciones alcanzan un nivel de similitud con el conjunto de criterios de selección de referencia.

La figura 2 presenta un esquema general del enfoque propuesto, cuyo proceso principal es el análisis de similitud de criterios de selección. Este análisis de similitud tiene como datos de entrada al conjunto de criterios explícitos, extraídos en forma manual a partir de documentos, tales como guías, estatutos, lineamientos, y otros documentos que definan los requerimientos de las posiciones de expertos. El análisis de similitud también toma como datos de entrada al conjunto de contenidos con criterios implícitos (descripciones), extraídos en forma automática a partir de las evaluaciones de los candidatos. El análisis de similitud extrae un conjunto de conceptos semánticos de cada conjunto de criterios (explícitos e implícitos), y calcula la frecuencia de los mismos. Finalmente, se determina el nivel de correlación de frecuencias de los conceptos semánticos correspondientes a cada grupo de criterios. El resultado final es un puntaje de similitud, asociado al nivel de correlación entre criterios explícitos e implícitos, o entre criterios implícitos.

2.1.1. Similitud semántica entre criterios de selección

El enfoque de elegibilidad de candidatos implica determinar la similitud entre conjuntos de criterios

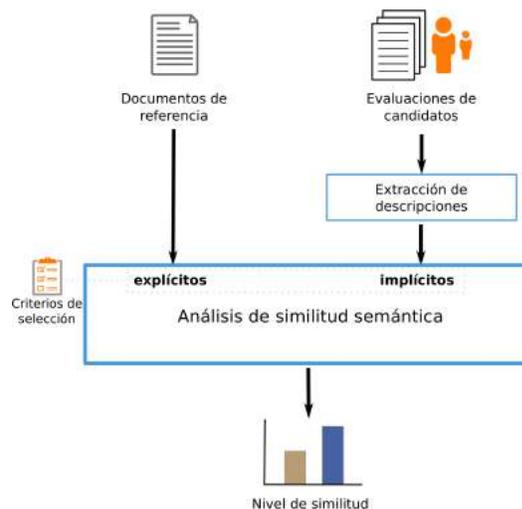


Figura 2: Enfoque general propuesto para el análisis de criterios de selección.

de selección mediante un análisis de las propiedades semánticas de éstos. El proceso propuesto incluye las siguientes etapas:

1. **Extraer criterios de selección contenidos en descripciones** de candidatos
Mediante el entrenamiento de un modelo de aprendizaje supervisado de clasificación binaria, cuyo objetivo es etiquetar sentencias en dos clases distintas: “descripciones” y “no descripciones”.
2. **Identificar las propiedades semánticas** de los criterios de selección
Mediante la aplicación de Etiquetado de Roles Semánticos (SRL)[4] a un conjunto de descripciones de candidatos, considerados contenedores de criterios de selección.
3. **Evaluar la relevancia de las propiedades**

semánticas de los criterios de selección

Mediante la determinación de relevancia de cada una de propiedades semánticas expresadas como la frecuencia de ocurrencia de los Roles Semánticos etiquetados en las descripciones.

4. **Evaluar la similitud entre criterios de selección**

Mediante el uso de un coeficiente de correlación de rankings sobre las frecuencias de las propiedades semánticas de los criterios de selección.

2.2. Selección de grupos de expertos mediante redes sociales

A diferencia de la selección de candidatos individuales, la selección de grupos no suele estar acompañada por evaluaciones de los grupos, que permitan analizar la aplicación de los criterios de selección. Sin embargo, una característica importante de la selección de grupos, es que los criterios de selección suelen considerar las relaciones entre los miembros del grupo. Por ejemplo, un grupo interdisciplinario con miembros “representativos” relacionaría los miembros con sus disciplinas, áreas geográficas, etc. La versatilidad de la información relacional representa una alternativa importante ante la falta de fuentes de información para la evaluación de criterios de selección de grupos de expertos.

Adicionalmente al problema de representación, la selección de grupos de expertos implica un elevado costo de procesamiento por la naturaleza combinatoria del problema, determinado por el coeficiente binomial ${}_nC_r$, con complejidad $O(n!)$.

En esta segunda parte del enfoque propuesto, se presenta un método para la selección de grupos de expertos, a partir del análisis de las redes sociales de un conjunto de candidatos elegibles. El método propuesto, utiliza un conjunto de métricas de SNA que permiten evaluar la conformidad de grupos a algunos criterios de selección de grupos. Asimismo, el método propone una estrategia de optimización para la búsqueda de conformaciones óptimas, mediante la implementación de un algoritmo genético.

2.2.1. Criterios de selección de grupos

El enfoque de selección de grupos de expertos presenta una estrategia para evaluar criterios de

selección de grupos en base a las relaciones sociales de sus miembros. Esta estrategia implica los siguientes aspectos principales:

1. **Representar información relacional** de los candidatos a conformar grupos de expertos

Mediante una estrategia de unificación de tipos de relaciones, para permitir la aplicación de varias de las métricas actuales de SNA.

2. **Determinar la independencia** entre los miembros de un grupo

A partir de una métrica de distancia social de dentro del grupo, como uno de los criterios más importantes para la selección de grupos de expertos.

La función para la métrica de grupos con miembros independientes queda definida como:

$$f = \frac{\left[\left(\sum_{i,j=0}^k d_{ij} \right) / k \right] + m}{2 * D} \quad (1)$$

donde d es la distancia geodésica entre dos miembros i y j del grupo, para $\forall i, j / i \neq j$ y $i, j \in N$, dado un conjunto de nodos N , donde k es el número de distancias entre miembros del grupo, D es el diámetro de la red, y m es la mínima distancia geodésica entre cada par de miembros del grupo.

3. **Integrar criterios** de selección de grupos basados en redes sociales

Mediante un enfoque de Toma de Decisiones con Múltiples Atributos (MADM).

El enfoque presenta la aplicación de una suma ponderada, junto con la identificación de pesos asignados a cada criterio mediante el método de Entropía [16].

La figura 3 presenta un esquema general de la propuesta de integración de criterios de selección de grupos.

4. **Optimizar el proceso de búsqueda**

Mediante la implementación de un Algoritmo Genético (GA).

El GA propuesto tiene por objetivo la búsqueda de conformaciones óptimas de grupos, mediante la evaluación de una función basada en la integración de métricas de SNA.

La representación de las soluciones utilizada para procesar el algoritmo genético es un vector de enteros que contiene el índice de nodos

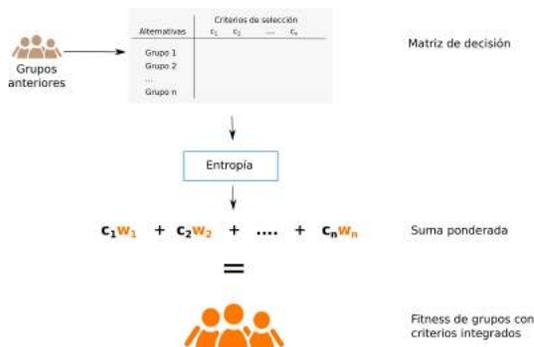


Figura 3: Esquema general de integración de criterios de selección con suma ponderada.

de la red social conformada por los candidatos, y un vector binario que contiene la información de inclusión en el grupo por parte de cada candidato.

3. Experimentos

Esta sección comienza con una descripción de los datos utilizados para la evaluación experimental del enfoque de elegibilidad de candidatos. Luego se presentan las actividades de clasificación de descripciones, y la conformidad de las evaluaciones respecto de los criterios de selección. Posteriormente, se presentan los datos utilizados para la evaluación experimental del enfoque de selección de grupos de expertos. Finalmente, se presentan los resultados de selección de grupos de expertos a partir de la métrica de independencia propuesta, y de la integración de criterios de selección basados en métricas de Análisis de Redes Sociales.

3.1. Datos para la evaluación de elegibilidad

Se desarrolló un nuevo conjunto de datos a partir de 65 documentos de texto correspondientes a audiencias de nominación del Committee of Commerce, Science, and Transportation del Congreso de los Estados Unidos. Estas nominaciones tuvieron lugar desde marzo de 2000 hasta julio de 2013¹. El conjunto de datos se constituyó de 7738 sentencias extraídas de las declaraciones de senadores, luego etiquetadas manualmente en clases

¹<http://www.gpo.gov/fdsysinfo/aboutfdsys.htm> (en línea en octubre de 2016)

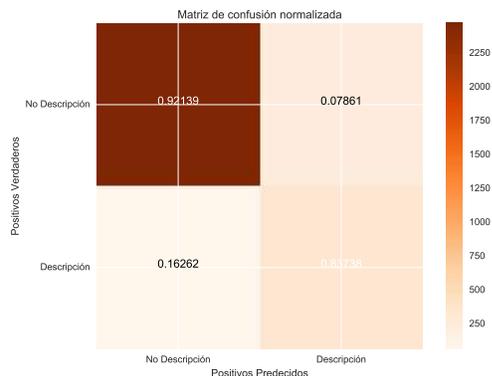


Figura 4: Matriz de confusión normalizada.

positivas (descripciones) y negativas (no descripciones).

Finalmente, el conjunto de datos presentó una distribución de clases asimétrica ($\approx 13\%$ positivas, $\approx 87\%$ negativas).

3.2. Clasificación de descripciones

El proceso de clasificación implicó la extracción de features (n-grams, $n=\{2,3,4\}$) dando 415176 n-gramas, y la selección de features mediante Chi cuadrado χ^2 (threshold=1000 features).

Se entrenó un modelo de aprendizaje supervisado mediante Support Vector Machines (SVM-RBF), y se optimizaron sus parámetros mediante Grid Search Cross Validation (GS-CV).

El tiempo total requerido para el entrenamiento del modelo demandó 270 segundos (4,5 minutos) aproximadamente en una computadora de escritorio estandar, compuesta por un procesador Intel I3 530 de 2.93GHz.

La figura 4 presenta la matriz de confusión normalizada.

El entrenamiento del modelo de clasificación propuesto alcanzó un desempeño de $f1\ score = 0,92$ y $ROC\ AUC = 0,88$ con los parámetros $C = 2$ y $gamma = 2e^3$.

La figura 5 presenta las curvas de aprendizaje (learning curves), las cuales presentan una reducción en el error al incrementar el tamaño del conjunto de datos de entrenamiento.

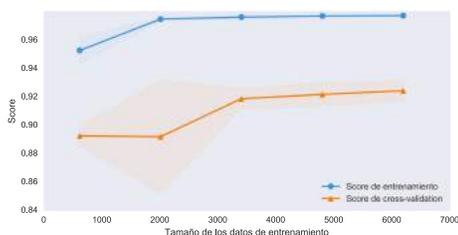


Figura 5: Curvas de aprendizaje del modelo entrenado para la clasificación de descripciones.

3.3. Conformidad de evaluaciones a los criterios de selección

Para determinar la conformidad de los criterios de selección se utilizaron dos conjuntos de datos, el de referencia asociado a los criterios explícitos, y el de correspondencia asociado a los criterios implícitos.

El conjunto de datos de criterios explícitos, se generó a partir de la información relacionada con las posiciones a cubrir por los candidatos. Mientras que el conjunto de datos de criterios implícitos, se generó a partir de las descripciones extraídas de las audiencias de nominación (ver cuadro 1).

Se evaluó la similitud entre los criterios implícitos en referencia a los criterios explícitos de tres segmentos: Comité, Departamento, y Directivo.

La figura 6 presenta los diagramas de caja de los coeficientes similitud para los segmentos Comité, Departamento, y Directivo. En la figura se observa un mejor desempeño en la similitud de los criterios de selección del segmento Departamento. La desviación estándar de los segmentos Comité ($\approx 0,35$) y Directivo ($\approx 0,05$) presentan una amplia diferencia. Sin embargo, debe considerarse la diferencia en la cantidad de nominaciones correspondientes a cada segmento (27 en Comité y 5 en Directivo).

3.4. Datos para la selección de grupos de expertos

Se implementó una red social de investigadores a partir de información pública extraída del sitio web de CONICET. Se aplicaron estrategias de unificación y desambiguación, ya que la mayor parte de la información disponible en el sitio web es ingresada por los mismos investigadores.

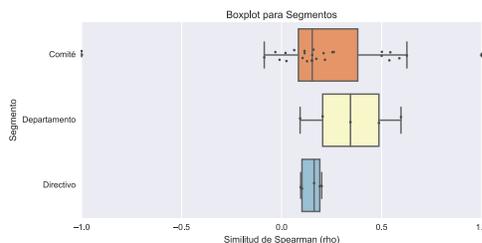


Figura 6: Diagramas de caja para los segmentos Comité, Departamento, y Directivo.

La elegibilidad de los candidatos, para la evaluación del enfoque propuesto de selección de grupos, se determinó a partir un listado de candidatos elegibles que cumplieran con los requisitos para conformar las comisiones. Sin embargo, se mantuvo la información de las relaciones con otros investigadores no elegibles, ya que todos ellos conforman la estructura de la red social.

Finalmente, la red social para el caso de estudio quedó conformada por un conjunto de 1293 nodos (investigadores) y 4322 enlaces (publicaciones y lugares de trabajo en común), conformando 74 componentes. Para la evaluación, se tomó el componente de mayor tamaño incluyó 1058 ($\approx 82\%$) de los investigadores (75 de ellos calificados para conformar comisiones), y 3878 ($\approx 90\%$) de relaciones.

3.5. Selección de grupos con miembros independientes

La primera evaluación experimental se realizó con la métrica de independencia de grupos como función de fitness del GA. Con objeto de comparar el fitness de las comisiones actuales con el fitness de las comisiones generadas por el algoritmo genético, se generaron comisiones de 3, 4, y 5 miembros.

La figura 7 presenta el fitness de las comisiones generadas y las actuales de acuerdo a las 16 configuraciones. En todas las ejecuciones se obtuvieron comisiones generadas con fitness superior al de las comisiones actuales.

La figura 8 muestra la red social utilizada en la experimentación, en la cual se identifica que los miembros de una comisión generada se encuentran más alejados entre sí que los miembros de la comisión actual. Esta representación muestra una mejora en el balance de las distancias entre los

Segmento	Fuentes para criterios explícitos	Fuentes para criterios implícitos
Comité	definiciones de criterios de selección del Federal Advisory Committee Act (FACA) ²	nominaciones a posición en mesas directivas o grupos consultivos
Departamento	definiciones de visión y misión de las dependencias correspondientes	nominaciones a dirección de sectores dentro de la misma dependencia
Directivo	definiciones de responsabilidades de cargos	nominaciones a posiciones dentro de una junta directiva

Cuadro 1: Orígenes de datos para la identificación de los criterios de selección explícitos (referencia) e implícitos (correspondencia) utilizados en la evaluación de similitud de criterios de selección aplicados.

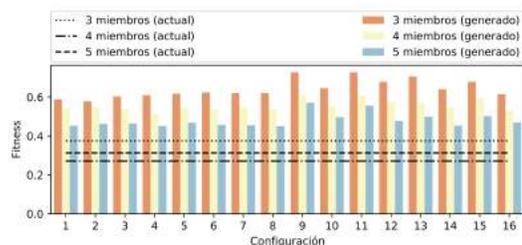


Figura 7: Fitness promedio para comisiones generadas de 3, 4, y 5 miembros.

membros de las comisiones generadas al comparlas con las comisiones actuales.

3.6. Selección de grupos con múltiples criterios

La segunda evaluación experimental se realizó mediante una suma ponderada de la métrica propuesta de Independencia de grupos y KPP-Pos[3]. Para ello, se determinaron los pesos correspondientes a cada métrica, mediante la aplicación del método de Entropía a los fitness de las comisiones actuales. Los pesos obtenidos para las métricas evaluadas indican una mayor relevancia dada a la selección de grupos con miembros independientes, dada por $w_{Independencia} = 0,567966$, que la relevancia dada a la selección de grupos que maximicen la difusión en la red, dada por $w_{KPP-Pos} = 0,432034$.

Al igual que en la primera evaluación, se comparó el fitness de las comisiones actuales con el fitness de las comisiones generadas por el algoritmo genético, por lo que se generaron comisiones de 3, 4, y 5 miembros.

Para la configuración del GA, se utilizó la configuración que arrojó mejores resultados en la primera evaluación. Esta configuración utilizó los me-

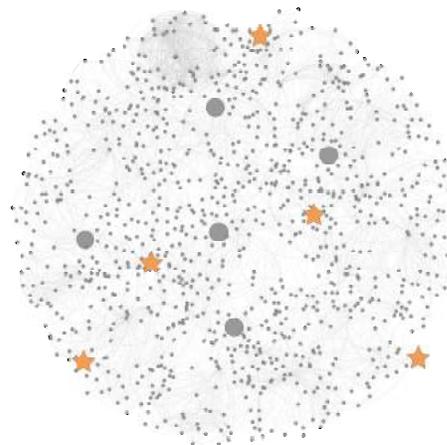


Figure 8: Comisión de 5 miembros. Actual (círculos grandes) y generada (estrellas).

canismos PMX, SWAP, SUS, y Generational.

Posteriormente se comparó el fitness de las comisiones generadas con 3, 4, y 5 miembros respecto del fitness de las comisiones actuales de Ingresos, Informes, y Becas del área de Informática y Comunicaciones, respectivamente.

La figura 9 presenta los fitness promedio, basado en la integración de las métricas KPP-pos e Independencia, de las comisiones generadas con 3, 4, y 5 miembros, junto con el fitness evaluado en las comisiones actuales. En la figura se observa que, al igual que en el caso de la evaluación de la métrica de Independencia en forma individual, al evaluar el fitness con las métricas integradas el método propuesto generó comisiones con mejores fitness que las comisiones actuales.

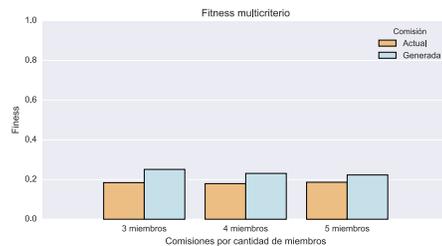


Figura 9: Fitness multicriterio de comisiones generadas y comisiones actuales.

4. Trabajos relacionados

El área de Recuperación de Expertos (RE), una subárea de Recuperación de Información (IR), define dos objetivos relacionados con el descubrimiento de asociaciones entre expertos y áreas de conocimiento [2]: el Perfilado de Expertos (Expert Profiling) y Hallazgo de Expertos (Expert Finding). Sin embargo, el objetivo de los sistemas de Hallazgo de Expertos es asociar expertos en un tema dado, mientras que el objetivo del Perfilado de Expertos es descubrir temas asociados a un experto.

Independientemente de su objetivo, los sistemas de RE necesitan encontrar temas o áreas de conocimiento mediante el análisis de evidencia que demuestre la experiencia del experto. En RE, esta evidencia de experiencia es generalmente representada por distintas fuentes de información como curriculum vitae, publicaciones científicas, y listas de áreas de conocimiento, entre otras.

La importancia de la evidencia de experiencia, así como el contexto, son requeridos para asistir a los usuarios de los sistemas de RE. En [1] se introduce la tarea de determinar los perfiles de expertos y proponen dos modelos para su resolución. Ambos modelos se basan en la construcción de perfiles basados en evidencia de experiencia tales como documentos y áreas de conocimiento.

La evidencia de experiencia es usualmente almacenada en documentos de texto en forma estructurada o semi estructurada. Por lo tanto, la extracción de evidencia de experiencia es uno de los principales desafíos de la RE. En [9] se reconoce la necesidad de transformar descripciones textuales de proyectos en un conjunto de habilidades que los candidatos expertos deben cumplir para ser seleccionados. Los autores proponen el concepto de *noción implícita de experiencia* en descripciones

de proyectos.

La mayoría de los enfoques para la extracción de evidencia de experiencia están basados en propiedades sintácticas de los documentos. En [10] se usan los curriculum vitae de un grupo de investigadores, como evidencia de experiencia en un modelo de recuperación de expertos basado en etiquetas. Los autores comparan el algoritmo utilizado con otros algoritmos para la recomendación de etiquetas que utilizan títulos, resúmenes, y palabras claves de contribuciones científicas. Otros trabajos se enfocan en la extracción de temas o tópicos, como representativos de áreas de conocimiento. En este sentido, en [14] se propone la extracción de temas de experiencia usando una adaptación de otros enfoques basados en extracción de términos y en extracción de frases clave. En [8] se argumenta que los enfoques actuales de RE, basados en la extracción de temas a partir de documentos, no consideran las relaciones ocultas entre los candidatos y las consultas. Con este fin, los autores proponen un enfoque para el Hallazgo de Expertos basado en Latent Dirichlet Allocation (LDA).

Otros trabajos en la extracción de experiencia tienen en cuenta otras propiedades de los documentos, además de las sintácticas o las semánticas. En [13] se aplica ML a la tarea de Perfilado de Expertos mediante el etiquetado automático de personas dentro de un contexto empresarial. Los autores proponen un listado de etiquetas con un orden de relevancia o ranking. Este enfoque utiliza un vocabulario de etiquetas creado por un conjunto de usuarios para este problema en particular.

Además de la evidencia de experiencia, otro aspecto importante sobre la creación de perfiles de expertos radica en la organización de los datos de dichos perfiles, así como de su inherente cambio. En [12] se introduce la tarea de identificar y caracterizar los cambios en la experiencia de los individuos en el tiempo. Los autores proponen el concepto de *perfil de experiencia jerárquica* para organizar la taxonomía de la experiencia, representada mediante un árbol ponderado. En esta línea, en [15] se considera la estructura jerárquica de las áreas de conocimiento relacionadas a un experto mediante un modelo de clasificación.

Algunos trabajos explotan particularidades del dominio de aplicación para la RE. En este sentido, en [7] se utilizan aspectos específicos del dominio de los sitios de pregunta-respuesta. Entre estos aspectos se incluyen la relevancia del tema, reputación del usuario, y la autoridad de una categoría.

En [6, 5] se modelan múltiples aspectos de la experiencia de personas en el marco del problema de Asignación de Revisores en Comités (CRA). Los autores proponen la extracción de temas a partir de documentos asociados a los candidatos expertos, como el caso de las contribuciones científicas.

En [11] se identifican un conjunto de criterios de selección, tanto incluyentes como excluyentes, de expertos para la conformación de paneles de asesores. El trabajo propone principios como la transparencia, disponibilidad, y público acceso de los criterios de elegibilidad y selección. Una diferencia fundamental con el trabajo propuesto es que la identificación de criterios se realiza sobre los lineamientos de las organizaciones relevadas, sin considerar la opinión de los responsables del proceso de selección. Por otra parte, los autores reconocen la importancia de la parcialidad (bias), y menciona algunas alternativas para su tratamiento, entre las que se encuentra el análisis de las declaraciones públicas de los candidatos, aunque no de las declaraciones de terceros.

5. Conclusiones

En este trabajo se propone un abordaje integral al problema de la selección de grupos de expertos basado en Aprendizaje Automático (ML) Procesamiento de Lenguaje Natural (NLP) y Análisis de Redes Sociales (SNA).

Con este trabajo, se pretende disponer de una estrategia para resolver la costosa tarea de analizar procesos de selección de expertos, y a su vez, sugerir conformaciones posibles de grupos de expertos en forma repetible y con pretensiones de objetividad. De esta manera, el enfoque propuesto es un aporte para dar claridad a los procesos de selección de grupos de expertos.

Los resultados obtenidos en la evaluación experimental indican que la correspondencia entre evaluaciones de los candidatos y los criterios de selección puede ser evaluada a partir de herramientas automatizadas de extracción y análisis de contenido semántico. Asimismo, la evaluación del método de selección de grupos de expertos basado en análisis de redes sociales, indican que es posible generar grupos con un mejor propiedades estructurales al compararlos con los comités actuales.

Independientemente de las contribuciones del trabajo, resulta pertinente indicar algunas consideraciones, como el costo del enfoque de aprendi-

zaje supervisado, el costo y estado actual de desarrollo de las alternativas de análisis semántico automático de textos, la representatividad del modelo de red social utilizado.

Actualmente, esta línea de investigación se continúa en el marco de una investigación posdoctoral, con especial interés en la construcción de perfiles de expertos en grandes conjuntos de datos, particularmente en el ámbito científico.

References

- [1] Balog, K., Azzopardi, L., & de Rijke, M. (2009). A language modeling framework for expert finding. *Information Processing & Management*, 45(1), 1–19.
- [2] Balog, K. & De Rijke, M. (2007). Determining Expert Profiles (With an Application to Expert Finding). In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, volume 7 of *IJCAI '07* (pp. 2657–2662).
- [3] Borgatti, S. P. (2006). Identifying sets of key players in a social network. *Computational & Mathematical Organization Theory*, 12(1), 21–34.
- [4] Gildea, D. & Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational linguistics*, 28(3), 245–288.
- [5] Karimzadehgan, M. & Zhai, C. (2012). Integer Linear Programming for Constrained Multi-Aspect Committee Review Assignment. *Information processing & management*, 48(4), 725–740.
- [6] Karimzadehgan, M., Zhai, C., & Belford, G. (2008). Multi-aspect expertise matching for review assignment. In *Proceedings of the 17th ACM conference on Information and knowledge management*, CIKM '08 (pp. 1113–1122).
- [7] Liu, D.-R., Chen, Y.-H., Kao, W.-C., & Wang, H.-W. (2013). Integrating expert profile, reputation and link analysis for expert finding in

- question-answering websites. *Information Processing & Management*, 49(1), 312–329.
- [8] Momtazi, S. & Naumann, F. (2013). Topic modeling for expert finding using latent Dirichlet allocation. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(5), 346–353.
- [9] Neshati, M., Beigy, H., & Hiemstra, D. (2014). Expert group formation using facility location analysis. *Information Processing & Management*, 50(2), 361–383.
- [10] Ribeiro, I. S., Santos, R. L., Gonçalves, M. A., & Laender, A. H. (2015). On tag recommendation for expertise profiling: A case study in the scientific domain. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, WSDM '15 (pp. 189–198).
- [11] Rowe, S., Alexander, N., Weaver, C. M., Dwyer, J. T., Drew, C., Applebaum, R. S., Atkinson, S., Clydesdale, F. M., Hentges, E., Higley, N. A., & Westring, M. E. (2013). How experts are chosen to inform public policy: Can the process be improved? *Health Policy*, 112(3), 172–178.
- [12] Rybak, J., Balog, K., & Nørsvåg, K. (2014). Temporal expertise profiling. In *European Conference on Information Retrieval*, ECIR '14 (pp. 540–546).
- [13] Serdyukov, P., Taylor, M., Vinay, V., Richardson, M., & White, R. W. (2011). Automatic people tagging for expertise profiling in the enterprise. In *European Conference on Information Retrieval*, ECIR '11 (pp. 399–410).
- [14] Silvello, G., Bordea, G., Ferro, N., Buitelaar, P., & Bogers, T. (2016). Semantic representation and enrichment of information retrieval experimental data. *International Journal on Digital Libraries*, (pp. 1–28).
- [15] Wu, T., Wang, Q., Zhang, Z., & Si, L. (2015). Determining expert research areas with multi-instance learning of hierarchical multi-label classification model. In *Proceedings of the 24th International Conference on Artificial Intelligence*, IJCAI '07 (pp. 2305–2511).
- [16] Xu, X. (2004). A note on the subjective and objective integrated approach to determine attribute weights. *European Journal of Operational Research*, 156(2), 530–532.
- [17] Zamudio, E., Berdun, L. S., & Amandi, A. (2013). Un algoritmo genético para la conformación de grupos de individuos distantes en redes sociales. In *Argentine Symposium on Artificial Intelligence*, 42 JAIIO Cordoba, Argentina.
- [18] Zamudio, E., Berdún, L. S., & Amandi, A. (2014). An approach to the creation of commissions of independent individuals using social networks and genetic algorithms. *Revista Iberoamericana de Inteligencia Artificial*, 17(53), 24–34.
- [19] Zamudio, E., Berdún, L. S., & Amandi, A. A. (2016). Social Networks and Genetic Algorithms to Choose Committees with Independent Members. *Expert Syst. Appl.*, 43(C), 261–270.

Reconocimiento de gestos dinámicos y su aplicación al lenguaje de señas

Franco Ronchetti

Directora en la UNLP: Laura Lanzarini

Director en la UTH-CUJAE: Alejandro Rosete

Fecha de exposición: 23/03/2017

Facultad de Informática, Universidad Nacional de La Plata

fronchetti@lidi.info.unlp.edu.ar

1. Introducción

El reconocimiento automático de gestos humanos es un problema multidisciplinar complejo y no resuelto aún de forma completa. Desde la aparición de tecnologías de captura de video digital existen intentos de reconocer gestos dinámicos con diferentes fines. La incorporación de nuevas tecnologías como sensores de profundidad o cámaras de alta resolución, así como la mayor capacidad de procesamiento de los dispositivos actuales, permiten el desarrollo de nuevas tecnologías capaces de detectar diferentes movimientos y actuar en tiempo real. A diferencia del reconocimiento de la voz hablada, que lleva más de 40 años de investigación, esta temática es relativamente nueva en el ambiente científico, y evoluciona de forma acelerada a medida que aparecen nuevos dispositivos, así como nuevos algoritmos de visión por computador.

La captura y reconocimiento de gestos dinámicos permite que sean utilizados en diversas áreas de aplicación como por ejemplo monitoreo de pacientes médicos, control en un entorno de videojuego, navegación y manipulación de entornos virtuales, traducción de léxicos de la lengua de señas, entre otras aplicaciones de interés. Particularmente la lengua de señas puede entenderse como un problema particular del reconocimiento de gestos dinámicos, el cual es sumamente apreciado en los últimos tiempos por distintas instituciones, ya que permite una ayuda directa a personas hipoacúsicas.

A grandes rasgos, se pueden distinguir gestos corporales, que se realizan con movimientos de todo el cuerpo, gestos con las manos, como un saludo, gestos con los dedos y las manos, como la lengua de señas y gestos faciales, como los guiños y movimientos de los labios (ver [Mit07]). Otra distinción importante es entre gestos estáticos, comúnmente llamados poses, definidos por una configuración particular del cuerpo en el entorno, y gestos dinámicos compuestos por una serie de movimientos de ciertas partes del cuerpo.

Actualmente, el uso de pantallas táctiles se ha convertido en un estándar para dispositivos móviles en ciertas aplicaciones; el reemplazo de los joysticks tradicionales por interfaces de voz y movimiento en las consolas de juegos se está consolidando. Sin embargo, el retiro del teclado-mouse en las PCs de propósito general por interfaces más naturales basadas en gestos, todavía se encuentra lejos de ser una realidad. En este panorama, las tecnologías más prometedoras para proveer una interfaz hombre-máquina eficiente son el reconocimiento de voz y de gestos en tiempo real [Kar06].

Para poder utilizar un sistema de reconocimiento automático de lengua de señas para traducir los gestos de un intérprete, es necesario afrontar una serie de diversas tareas. En primer lugar, existen diferentes enfoques dependiendo el dispositivo de sensado a utilizar. Si bien existen dispositivos invasivos como guantes de datos, en esta Tesis se analizan sólo dispositivos no invasivos de dos tipos: las cámaras RGB convencionales, y las cámaras de profundidad (con particular interés en los nuevos dispositivos RGB-d). Una vez capturado el gesto se requiere de diversas etapas de pre-procesamiento

para identificar regiones de interés como las manos y rostro del sujeto/intérprete, para luego identificar las diferentes trayectorias del gesto realizado. Además, particularmente para la lengua de señas existe una variabilidad enorme en las diferentes posturas o configuraciones que la mano puede tener, lo cual hace a esta disciplina una problemática particularmente compleja. Para afrontar esto es necesario una correcta generación de descriptores tanto estáticos como dinámicos. Este es uno de los ejes principales investigados en esta Tesis.

Capturar, analizar y responder ante un evento es una tarea sumamente compleja que involucra diferentes áreas de la informática como:

- Procesamiento de imágenes. En todo proceso de reconocimiento en video es necesario contar con un adecuado manejo y filtrado de imágenes. Esto puede involucrar, entre otras cosas, eliminación de ruido, escalado/rotado de la imagen, filtros frecuenciales para detección de patrones, filtros de color, etc.
- Procesamiento temporal. Ya que se entiende un gesto como una secuencia de movimiento de una o varias partes del cuerpo, es necesario realizar un adecuado procesamiento de la información temporal.
- Sistemas Inteligentes. Para realizar la clasificación de un patrón de video y poder actuar en consecuencia, es necesario la utilización de técnicas inteligentes (*machine learning*).

Las lenguas de señas pueden entenderse como un caso particular del reconocimiento de gestos dinámicos. Es un problema multidisciplinar sumamente complejo con muchas aristas a mejorar en la actualidad. Si bien recientemente ha habido algunos avances a través del reconocimiento de gestos, todavía hay un largo camino por recorrer antes de poder tener aplicaciones precisas y robustas que permiten traducir e interpretar los signos realizados por un intérprete [Coo11].

La tarea de reconocer una lengua de señas implica un proceso de múltiples pasos, que puede ser simplificado del siguiente modo:

1. El seguimiento de las manos del intérprete
2. La segmentación de las manos y la creación de un modelo de su forma
3. Reconocimiento de las formas de las manos
4. Reconocimiento del signo como una entidad sintáctica
5. Asignación de semántica a una secuencia de signos
6. Traducción de la semántica de los signos a la lengua escrita

Estos pasos se detallan en el capítulo 2. Si bien estas tareas pueden proporcionar información entre ellas, de modo general pueden ser llevadas a cabo independientemente, y de diferentes maneras. Por ejemplo, hay varios enfoques para el seguimiento de movimientos de la mano: algunos utilizan sistemas 3D [Sou14, Pug11], tales como MS Kinect. Otros simplemente utilizan una imagen 2D a partir de una cámara [Coo11, Von08]. La mayoría de los sistemas más antiguos emplean sensores de movimiento tales como guantes especiales, acelerómetros, etc., aunque los enfoques más recientes se centran generalmente en el procesamiento de vídeo. Existen numerosas publicaciones sobre el reconocimiento automático de las lenguas de señas, un campo que comenzó hacia los años 90. Puede verse en [Kol15], [Von08] y [Coo11] algunas revisiones generales del estado del arte en esta temática.

El reconocimiento del lenguaje de señas emplea diferentes tipos de características, generalmente clasificadas como manuales y no manuales. Las características no manuales, como pueden ser la postura, lectura de labios o la cara del intérprete se incluyen a veces para mejorar el proceso de reconocimiento, ya que algunas señales no pueden ser diferenciadas únicamente con información manual [Von08]. En este sentido, por ejemplo, el seguimiento de la cabeza es un problema mayormente resuelto [Vio04], pero su segmentación con respecto a un fondo arbitrario o

en presencia de oclusiones mano-cabeza sigue siendo un problema sin resolver. No obstante, la información manual suele contener la mayor parte de la información en una señal.

Para el seguimiento y la segmentación de las manos, hay mucho interés en la creación de modelos de color de la piel para detectar y realizar un seguimiento de las manos de un intérprete en un video [Rou10], y añadiendo la posibilidad de segmentar las manos [Coo12], incluso en presencia de oclusiones mano-mano [Zie05].

La información de la configuración de una mano de un signo está compuesta por una secuencia de poses de esa mano [Von08]. Luego de la segmentación, la mano debe ser representada en una forma conveniente para el reconocimiento de la configuración. Conseguir una representación adecuada fotograma a fotograma de la mano no es una tarea trivial, existiendo diferentes estrategias que aproximan a una solución óptima. Mientras que la mejor salida posible a partir de este paso sería un modelo completo en 3D de la mano, esto es generalmente difícil de hacer sin múltiples cámaras, sensores o marcadores especiales [Pug11]. En la mayoría de los casos, la configuración de la mano en su lugar se representa como una combinación de características más abstractas basada en propiedades geométricas o morfológicas de su forma o textura [Von08].

Algunos investigadores se centran en el reconocimiento dactilológico (*fingerspelling*) [Pug11], que es esencialmente una tarea de reconocimiento de configuración estática. Mientras que algunos signos de hecho presentan una configuración de la mano estática en una o ambas manos, y no hay movimiento, la mayoría implican muchas formas manuales y sus transiciones, o transformaciones de una sola forma de la mano (rotación y traslación, etc.), y un cierto movimiento de las manos. Para hacer frente a estas señales dinámicas, los sistemas de reconocimiento de gestos (SLR, *Sign Language recognition*) generalmente se basan en Modelos Ocultos de Markov (HMMs), Deformación Dinámica de Tiempo (DTW) o modelos similares, ya sea para reconocer las señales segmentadas o un *stream* continuo ([Von08, Coo11]).

2. Objetivos

El objetivo general de esta tesis es desarrollar un modelo de reconocimiento automático de la Lengua de Señas Argentina (LSA). Esto trae aparejados los siguientes objetivos específicos:

- Analizar, describir y comparar las diferentes estrategias existentes en el estado del arte sobre reconocimiento y segmentación de manos.
- Construir una base de datos con fotografías de configuraciones de manos de la LSA utilizando marcadores de color para simplificar la segmentación.
- Realizar un método de clasificación de configuraciones de manos incluyendo la adecuada generación de descriptores.
- Construir una base de datos de la LSA con gestos dinámicos que permitan tanto la implementación de traductores específicos para la región, así como también dar la posibilidad a otros investigadores de poder utilizar el repositorio como herramienta de pruebas para algoritmos de aprendizaje automático.
- Realizar un método de clasificación modular de señas segmentadas que permita reconocer diferentes gestos, así como la posibilidad de intercambiar partes del clasificador para evaluar distintos métodos.

3. Contribuciones

Las principales contribuciones de esta tesis son las siguientes:

- Una revisión bibliográfica actualizada sobre diferentes estrategias de clasificación de gestos estáticos y dinámicos, incluyendo descriptores de imágenes, video y algoritmos inteligentes de clasificación, así como una revisión de las bases de datos de gestos existentes en la literatura.
- Dos bases de datos de la Lengua de Señas Argentina inexistentes hasta el momento. LSA16 contiene fotografías de 10 individuos distintos para 16 configuraciones de manos de las más utilizadas en el léxico argentino, con un total de 800 imágenes correctamente etiquetadas para cualquier proceso de aprendizaje automático. LSA64 es una base de datos de señas capturadas con una cámara de video de alta resolución. Contiene 64 señas distintas del LSA interpretadas por 10 sujetos distintos con un total de 3200 videos, correctamente etiquetados y con una versión preprocesada donde se tiene información del seguimiento y segmentación de las manos.
- Un método de clasificación de configuraciones del lenguaje de señas, junto con un conjunto de descriptores que posibilitan reconocer las diferentes formas que puede tener las manos de un intérprete. Este método puede utilizarse tanto para el lenguaje de señas como para cualquier aplicación donde se requiera identificar diferentes posturas de las manos.
- Un método probabilístico para clasificar señas basado en tres componentes principales: la posición, la configuración, y el movimiento de cada mano. Este método, basado en componentes posibilita el análisis de cada módulo por separado, dando la posibilidad de intercambiar sub-clasificadores por otros. El modelo propone un análisis específico de la información, creando descriptores apropiados para cada módulo, junto con métodos de clasificación independientes.

4. Base de datos de la Lengua de Señas Argentina (LSA)

En el ámbito de la lengua de señas existen diversas bases de datos dependiendo del problema al cual se dirigen. Aquí se distinguen tres tipos principales: reconocimiento de configuraciones de manos, reconocimiento de una seña, y reconocimiento de una sentencia. Cada tipo de base de datos presenta un desafío mayor que el anterior y permite experimentar con mayores pasos en las etapas de reconocimiento [Von08, Co011].

Las bases de datos aquí presentadas tienen dos objetivos específicos: por un lado, al estar grabadas con guantes de color en las manos de los intérpretes, permiten la rápida segmentación y seguimiento de las manos, utilizando únicamente un accesorio simple, barato y fácil de conseguir. Por otro lado, el conjunto de datos pretende ser un primer paso en la construcción de una base de datos completa para el léxico argentino, inexistente hasta el momento.

Un aspecto esencial de las lenguas de señas es el tipo de configuración que la seña posee. En ocasiones, diversas señas sólo se diferencian por la configuración de la mano, siendo el movimiento que se realiza el mismo que en otras. Esto lleva a la necesidad de tener como parte de un reconocimiento de señas, una etapa de clasificación de configuraciones de manos. Incluso, es normal que una sea comienza con una configuración y termine con otra, o con variaciones de la misma.

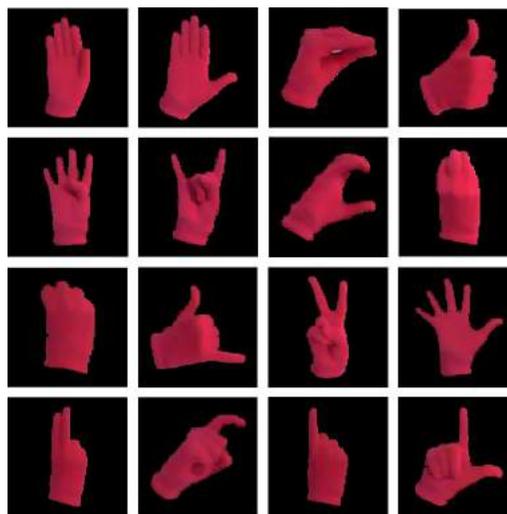


Figura 1. Las 16 configuraciones de manos en la base de datos LSA16

La primera base de datos creada, llamada LSA16, contiene 800 imágenes en donde 10 sujetos realizaron 5 repeticiones de 16 tipos distintos de configuraciones de manos utilizadas en distintas señas del léxico. La figura 1 muestra un ejemplo de cada una de las 16 clases dentro de la base de datos. Las imágenes son la segmentación de las fotografías reales. Cada configuración fue realizada repetidamente en diferentes posiciones y diferentes rotaciones en el plano perpendicular a la cámara, para generar mayor diversidad y realismo en la base de datos. Se utilizó una cámara web Logitech con 640x480 píxeles.

Los sujetos vistieron ropa negra, sobre un fondo blanco con iluminación controlada. Para la simplificar el problema de segmentación de la mano dentro de una imagen, los sujetos utilizaron guantes de tela con colores fluorescentes en sus manos. Esto resuelve parcialmente pero de un modo muy eficaz el reconocimiento de la posición de la mano y carece de los problemas existentes en los modelos de piel. Por otro lado, propone un artefacto simple y económico al momento de realizar pruebas o desarrollar una aplicación real. Los detalles de esta base de datos pueden encontrarse publicados en [Ron16c].

Como se mencionó anteriormente, cada región en el mundo tiene su léxico particular en lengua de señas. Esto hace imposible utilizar una base de datos extranjera si se quiere desarrollar un traductor argentino. Por otro lado, se estableció también que debido a la complejidad que se requiere para segmentar las manos de los intérpretes, las bases de datos actuales resultan difíciles de abordar para evaluar la eficacia de un modelo de clasificación.

En segundo lugar, se creó la base de datos LSA64, primer conjunto de videos específico para la Lengua de Señas Argentina. Este conjunto de datos contiene un total de 3200 videos en formato FullHD con 60FPS de 10 sujetos distintos interpretando 64 señas del LSA. La base de datos está públicamente disponible junto con una versión pre-procesada de la misma, para facilitar a los investigadores algunos pasos de segmentación. Cada intérprete utilizó dos guantes de color diferente con el fin de realizar la tarea de segmentación de forma rápida y eficiente. Esta estrategia puede verse utilizada en trabajos anterior, como por ejemplo en [Wan09], donde se utilizaron guantes no sólo para segmentar la mano sino para facilitar la clasificación de la configuración. La figura 2 muestra dos ejemplos tomados de la base de datos.

La base de datos fue construida en dos sets de grabación distintos. En el primero fueron grabadas 23 señas con una sola mano y se utilizó luz natural en un entorno abierto. En el segundo set, se agregaron 41 señas más (22 con dos manos, y 19 con una mano) y se utilizó luz artificial en un entorno semi-cerrado. Estas diferencias en iluminación permiten entrenar un modelo robusto que funcione en diferentes entornos.

Si bien 64 señas no es un número particularmente grande para los léxicos reales de lenguas de señas, es un paso inicial para construcción de una base de datos más robusta del léxico argentino, al mismo tiempo que permite un desafío para cualquier sistema de reconocimiento de gestos dinámicos. Las 64 señas elegidas poseen una gran diversidad, presentando superposición tanto de movimientos como en configuración de las manos, siendo necesario analizar todos los aspectos que la componen. Al mismo tiempo los 10 sujetos distintos existentes en la base posibilitan el estudio de un sistema no dependiente al sujeto. Los detalles de esta base de datos pueden encontrarse publicados en [Ron16b].



Figura 2. Dos fotogramas de ejemplos de la base de datos LSA64.

5. Modelo de Clasificación propuesto

El modelo desarrollado en esta tesis consta de diversas etapas para lograr la clasificación de gestos segmentados. La figura 3 muestra un esquema general del modelo donde pueden observarse los detalles de cada etapa. El esquema de clasificación se basa en un sistema probabilístico que tiene en cuenta la información de ambas manos. Si bien el foco del trabajo está puesto en la clasificación de lengua de señas, es posible adaptar el modelo a cualquier tipo de gesto corporal. De cada mano se utilizan tres componentes esenciales en una seña: la posición, la configuración, y el movimiento de la mano. Para obtener las probabilidades para cada uno de estos componentes se definieron diferentes clasificadores parciales, abordando las características que cada problema presenta. Los detalles del modelo aquí expuesto se encuentran publicados en [Ron16a] y [Ron15].

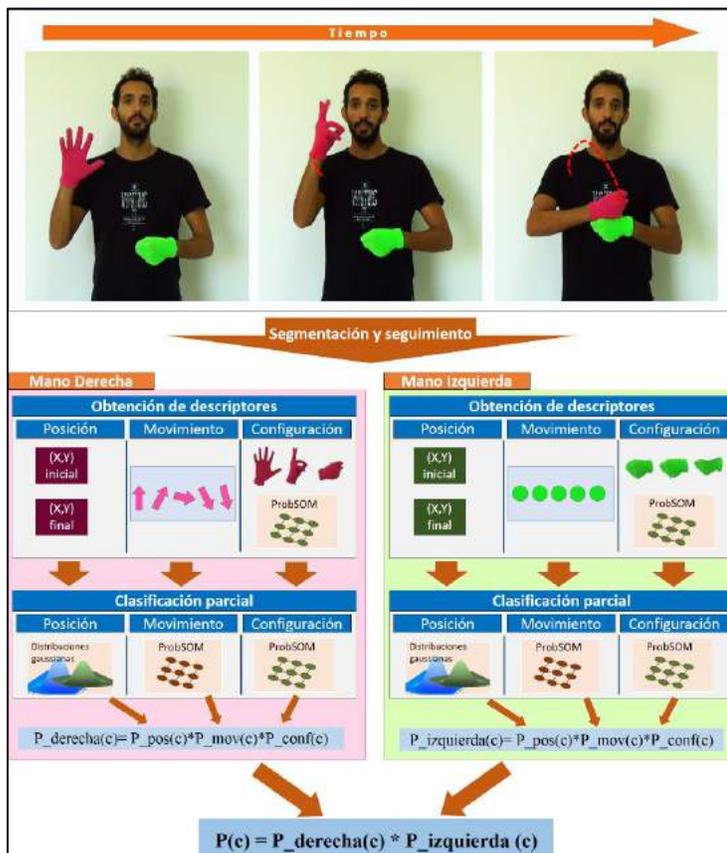


Figura 3. Descripción general del modelo de clasificación propuesto para señas segmentadas

El proceso de clasificación de una seña desde el video segmentado puede simplificarse en los siguientes pasos:

1. **Segmentación y Seguimiento.** Inicialmente es necesaria una etapa de segmentación donde cada mano es identificada a partir de un filtro de color. Aquí se obtiene no sólo la posición sino también una máscara con la forma de la mano, lo que facilita luego la clasificación de la configuración. El seguimiento de las manos se realiza fotograma a fotograma almacenando la posición de cada mano relativa a la posición de la cabeza del intérprete.
2. **Generación de descriptores.** En segundo lugar, es necesario generar descriptores apropiados para reconocer los tres componentes principales de la seña. Para describir la posición de cada mano se utilizaron las coordenadas 2D del primer y último fotograma del video. Para describir el movimiento se utilizaron diferencias de percentiles de las posiciones de cada mano. Por último, para describir la configuración de cada mano se utilizó un modelo de clasificación basado en el ProbSOM [Est10].
3. **Clasificación parcial.** En tercer lugar, se realiza un proceso de clasificación parcial. Dado cada conjunto de descriptores, se clasifican de forma independiente, obteniendo una probabilidad parcial de pertenencia a cada clase. Para clasificar la posición de cada mano se utilizó un sistema de distribución de gaussianas, considerando el conjunto de las diferentes posiciones como una distribución normal. Tanto para la clasificación del movimiento como de la configuración de manos se utilizaron redes ProbSOM.
4. **Clasificación de una seña.** Por último, los resultados de los clasificadores parciales son utilizados como entrada para el clasificador probabilístico total. La probabilidad de que una

seña pertenezca a una clase se computa como el producto de probabilidades de cada mano calculada de forma independiente. A su vez, la probabilidad de cada mano se calcula como el producto de probabilidades obtenidas de los clasificadores de posición, configuración y movimiento.

6. Trabajos experimentales

Con el fin de validar el modelo desarrollado, se realizaron diversas etapas de experimentación en las bases de datos desarrolladas, como así también en una base de datos de gestos dinámicos capturados con el dispositivo MS Kinect [Ron15]. Todos los resultados aquí mostraron son el promedio de realizar un proceso de validación cruzada aleatoria, con 30 pruebas independientes, 90% de datos para entrenamiento, y 10% para validación.

La tabla 1 muestra los resultados obtenidos de los experimentos realizados en la base de datos LSA16. Se evaluaron los clasificadores ProbSOM, así como también métodos clásicos como Máquinas de Soporte Vectorial (SVM) y Redes Neuronales Feedforward. Dos tipos de descriptores fueron computados sobre las imágenes: vectores SIFT y la Transformada de Radón. Luego, utilizando la mejor configuración obtenida (descriptor Radon y ProbSOM) se llevó a cabo una validación cruzada inter-sujeto, dejando un sujeto para validación y entrenando con el resto. La media de los 10 sujetos con $n = 30$ repeticiones independientes fue de 87,9%(±4,7%). Como es de esperar, al dejar un sujeto fuera, la tasa de acierto decae, ya que cada persona realiza las configuraciones de forma particular, con tamaños y apariencia de mano propia del individuo. No obstante, el sistema sigue mostrando buenos resultados, dando como posibilidad el reconocimiento correcto de una configuración realizada por un nuevo individuo desconocido por el sistema.

Método	Precisión
ProbSom con Radon	92,3(±2,05)
ProbSom con SIFT	88,7(±2,50)
Random Forest con Radon	91,0(±1,91)
SVM con Radon	91,2(±1,69)
Red Neuronal Feedforward con Radon	78,8(±3,80)

Tabla 1. Precisión del modelo para la base de datos LSA16 de configuraciones.

Para validar el modelo de clasificación de señas segmentadas, se realizaron una serie de evaluaciones sobre la base de datos LSA64 utilizando el modelo

	Todos	Config	Mov	Pos	Config-Pos	Config-Mov	Pos-Mov	Todos-HMM	Todos-BF-SVM
μ	97.44	52.97	54.03	76.05	94.91	83.59	84.84	95.92	95.08
σ	0.59	1.74	1.71	0.62	0.52	0.87	0.90	0.95	0.69

Tabla 2. Resultados de los experimentos llevados a cabo sobre la base de datos LSA64

propuesto y comparando los resultados al quitar parte de los sub-clasificadores propuestos. Las diferentes pruebas llevadas a cabo muestran la importancia de cada componente, ya que al quitar alguno la tasa de acierto decae, mostrando que cada componente agrega información no redundante al sistema. Por otro lado, se realizaron evaluaciones de comparación tanto para los descriptores como para el clasificador. Por un lado, la columna Todos-HMM muestra los resultados al reemplazar los subclasificadores de trayectoria y configuración por Modelos Ocultos de Markov (HMM) con Modelos de Mixturas Gaussianas (GMM). El resultado aquí fue de casi 96%. Esto muestra que si bien el ProbSOM obtuvo una mejora al reducir el error en un 60% comparado a los Modelos Ocultos de Markov, los descriptores propuestos son una parte esencial en el proceso de clasificación. Por otro lado, la última columna, titulada Todos-BF-SVM, muestra los resultados al cambiar la obtención de descriptores por los *binary features* propuestos por Kadir en [Kad04], clasificador con una Máquina de Soporte Vectorial.

Otro aspecto importante a considerar en la evaluación de los resultados es si la seña se realizó con una o dos manos. Si bien el modelo se definió para evaluar ambas manos, la base de datos utilizada (LSA64) posee tanto señas con dos manos, como señas con una sola mano (la dominante). En el primer caso, al tener ambas manos, podría conllevar a una ventaja en la clasificación, ya que se está incorporando información al modelo. En este sentido, es importante evaluar cómo se comporta el modelo en ambos casos. Siguiendo esta idea, se dividió la base de datos en dos subconjuntos de datos, uno con las 22 clases de señas con dos manos, y otro con las 42 clases de señas con una mano. Se realizaron experimentos independientes para verificar la tasa de acierto del modelo. Mientras que la media entre ambos tipos de experimentos no difiere significativamente de los resultados de las pruebas generales, cabe rescatar el aumento de la tasa de acierto al tener sólo clases de dos manos, logrando un error de sólo 0,91%. Cabe rescatar que la tasa de acierto al tener señas de sólo una mano llega a casi 96%, mostrando excelentes resultados para las 42 señas con esta característica.

Por último, y al igual que para LSA16, se llevaron a cabo una serie de experimentos sobre LSA64 para analizar

Sujeto	1	2	3	4	5	6	7	8	9	10	Media
μ	94.5	93.8	87.7	93.8	91.8	92.6	89.1	90.3	88.4	94.6	91.7
σ	0.66	0.83	1.05	0.79	0.65	0.41	0.91	0.70	0.85	0.66	0.75

Tabla 3. Validación independiente al sujeto sobre LSA64.

cómo se comporta el sistema ante la presencia de un sujeto nuevo, con el que no se había entrenado previamente. Para esto se entrenó el sistema con 9 sujetos, dejando el restante para validación, haciendo una evaluación cruzada inter-sujeto con 30 ejecuciones independientes. La tabla 3 muestra los resultados obtenidos para cada uno de los 10 sujetos de la base de datos LSA64, junto con la media de todos los resultados. Como es de esperar, la tasa de acierto decae con respecto a los resultados generales. No obstante, la precisión media conseguida fue de 91,7%(±0,8), mostrando excelentes resultados al introducir un nuevo individuo al sistema.

7. Conclusiones y líneas de trabajo futuras

La Tesis aquí resumida cuenta con dos aportes principales: por un lado, un modelo de clasificación para gestos dinámicos específicamente diseñado para la lengua de señas. Por otro lado, una base de datos multimedia de la Lengua de Señas Argentina, inexistente hasta el momento.

La base de datos desarrollada, llamada LSA64, posee 64 señas distintas del LSA. Los intérpretes utilizaron guantes de color para facilitar la segmentación de las manos. Este proceso resulta accesible para cualquiera ya que el guante es una herramienta económica y de fácil acceso. La base de datos propone tanto un diccionario específico para el léxico argentino como una herramienta de base de pruebas para cualquier trabajo de aprendizaje automático. La base de datos consta de 10 sujetos interpretando cada seña 5 repeticiones distintas, dando un total de 3200 videos de alta resolución. Sumado a esto, la base de datos LSA16 contiene 800 imágenes de configuraciones de manos del léxico argentino.

El método propuesto para clasificación de señas propone un esquema modular con subclasificadores parciales capaces de interpretar tres características principales en una seña: la posición, el movimiento y la configuración. Como sub-clasificador de configuración se utilizó también una red tipo PromSOM para clasificar las 16 configuraciones del LSA. Este trabajo fue primero evaluado por separado para general descriptores apropiados que permitieron luego adicionar la información temporal de las diferentes configuraciones que puede tener una seña. Por último, como sub-clasificadores de las posiciones de las señas, se utilizaron distribuciones estadísticas con modelos gaussianos de las posiciones iniciales y finales que cada mano posee en una seña.

Los resultados obtenidos sobre la clasificación de configuraciones mostraron ser relevantes y factibles de utilizar en un entorno real. Además, siendo un clasificador probabilístico, el sistema posee la capacidad de poder ser utilizado como descriptor para el clasificador parcial del modelo propuesto. Los experimentos realizados sobre la base de datos LSA64 fueron extensos y con resultados satisfactorios. Se realizaron diferentes pruebas sobre los clasificadores parciales para observar su comportamiento al igual que diferentes evaluaciones sobre el clasificador completo demostrando su robustez ante diferentes escenarios como por ejemplo la incorporación de un nuevo sujeto al sistema.

Existen diversas líneas de investigación que quedan abiertas luego de la finalización de esta tesis, entre las que se cabe nombrar:

- Focalizarse en la etapa de detección de manos para poder realizar una segmentación sin necesidad de marcadores de color. Esto permitiría realizar pruebas en otras bases de datos existentes donde no existe información de las posiciones de las manos ni tampoco se utilizan estos tipos de marcadores. Una de las estrategias más recientes aplicadas a esta temática son las redes convolucionales, relacionadas con el concepto de aprendizaje profundo (*deep learning*), que recién está emergiendo.
- Si bien se utilizaron algunos descriptores y clasificadores propuestos por otros autores en el estado del arte, generalmente esta tarea resulta sumamente compleja debido a que el modo de obtener los descriptores de una seña está relacionado con el clasificador propuesto.
- Para poder llevar a cabo un traductor más robusto sin duda es necesario aumentar el número de señas en la base de datos. Esto supone un desafío importante, no sólo por el tiempo y puesta en escena de las grabaciones requeridas sino también porque aumentar considerablemente el número de clases en una base de datos implica aumentar el error de clasificación en cualquier proceso de aprendizaje automático.
- Incorporar información no manual, como pueden ser expresiones de la cara, lectura de labios, inclinación del torso, etc. Este es un trabajo que algunos investigadores ya están abordando. El lenguaje de señas no sólo se basa en los movimientos de las manos sino en realizar un diccionario completo involucra también evaluar información no-manual, principalmente del rostro.

8. Referencias

- Coo11 Helen Cooper, Brian Holt, and Richard Bowden. Sign language recognition. In Thomas B. Moeslund, Adrian Hilton, Volker Krüger, and Leonid Sigal, editors, *Visual Analysis of Humans: Looking at People*, chapter 27, pages 539 – 562. Springer, 2011.
- Coo12 Helen Cooper, Eng-Jon Ong, Nicolas Pugeault, and Richard Bowden. Sign language recognition using sub-units. *Journal of Machine Learning Research*, 13:2205–2231, Jul 2012.
- Est10 Cesar Estrebow, Laura Lanzarini, and Waldo Hasperue. Voice recognition based on probabilistic som. In *Latinamerican Informatics Conference. CLEI 2010*. Paraguay. October 2010.
- Kad04 T. Kadir, R. Bowden, Ej Ong, and a. Zisserman. Minimal Training, Large Lexicon, Unconstrained Sign Language Recognition. *British Machine Vision Conference*, pages 96.1–96.10, 2004.
- Kar06 Maria Karam. PhD Thesis: A framework for research and design of gesturebased human-computer interactions. PhD thesis, University of Southampton, October 2006.

- Kol15 Oscar Koller, Jens Forster, and Hermann Ney. Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers. *Computer Vision and Image Understanding*, 141:108–125, December 2015.
- Mit07 S. Mitra and T. Acharya. Gesture recognition: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37(3):311–324, 2007.
- Pug11 N. Pugeault and R. Bowden. Spelling it out: Real-time ASL fingerspelling recognition. In *1st IEEE Workshop on Consumer Depth Cameras for Computer Vision*, in conjunction with ICCV'2011, 2011.
- Ron15 Ronchetti, Facundo Quiroga, Laura Lanzarini, César Estrebou. Distribution of Action Movements (DAM): A Descriptor for Human Action Recognition. *Franco Frontiers of Computer Science*. ISSN 2095-2236. Springer, Higher Education Press. v9. pp956-965. Diciembre 2015.
- Ron16a Franco Ronchetti, Facundo Quiroga, César Estrebou, Laura Lanzarini, Alejandro Rosete. Sign Language Recognition without frame-sequencing constraints: A proof of concept on the Argentinian Sign Language. *Advances in Artificial Intelligence - IBERAMIA 2016: 15th Ibero-American Conference on AI*, San José, Costa Rica, November 23-25, 2016, Proceedings. pp338-349. Springer International Publishing. 2016.
- Ron16b Franco Ronchetti, Facundo Quiroga, César Estrebou, Laura Lanzarini, Alejandro Rosete. LSA64: An Argentinian Sign Language Dataset. *XXII Congreso Argentino de Ciencias de la Computación. CACIC 2016*. San Luis. Argentina. pp794-803. Octubre 2016.
- Ron16c Franco Ronchetti, Facundo Quiroga, César Estrebou, Laura Lanzarini. Handshake recognition for Argentinian Sign Language using ProbSom. *Journal of Computer Science & Technology*. ISSN 1666-6038. Editorial ISTECS – RedUNCI. 16:1, pp01-05. April 2016.
- Rou10 Anastasios Roussos, Stavros Theodorakis, Vassilis Pitsikalis, and Petros Maragos. Hand tracking and affine shape-appearance handshake sub-units in continuous sign language recognition. In *Trends and Topics in Computer Vision - ECCV 2010 Workshops*, Heraklion, Crete, Greece, September 10-11, 2010, Revised Selected Papers, Part I, pages 258–272, 2010.
- Sou14 Gabriel de Souza Pereira Moreira, Gustavo Ravanhani Matuck, Osamu Saotome, and Adilson Marques da Cunha. Recognizing the brazilian signs language alphabet with neural networks over visual 3d data sensor. In *Advances in Artificial Intelligence-IBERAMIA 2014*, pages 637–648. Springer, 2014.
- Vio04 Paul Viola and Michael J Jones. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004.
- Von08 Ulrich von Agris, Jörg Zieren, Ulrich Canzler, Britta Bauer, and Karl-Friedrich Kraiss. Recent developments in visual sign language recognition. *Universal Access in the Information Society*, 6(4):323–362, 2008.
- Wan09 Robert Y. Wang and Jovan Popovic. Real-time hand-tracking with a color glove. *ACM Transactions on Graphics*, 28(3), 2009.
- Zie05 Jörg Zieren and Karl-Friedrich Kraiss. Robust person-independent visual sign language recognition. In *Pattern recognition and image analysis*, pages 520–528. Springer, 2005.

Formalización y Generalización del Manejo de Preferencias en Servicios de Razonamiento Rebatible

Juan Carlos Lionel Teze

Directores de tesis: **Guillermo R. Simari y Alejandro J. García**

Fecha de exposición: 30/03/2017

Instituto de Ciencias e Ingeniería de la Computación (ICIC),
Dep. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur
Laboratorio de Investigación y Desarrollo en Inteligencia Artificial Concordia,
Facultad de Ciencias de la Administración, Universidad Nacional de Entre Ríos
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Av. Alem 1253, (B8000CPB) Bahía Blanca, Argentina
Tel: (0291) 459-5135 / Fax: (0291) 459-5136
email:jct@cs.uns.edu.ar

Resumen

En los últimos años, la argumentación rebatible ha realizado un importante aporte a la Inteligencia Artificial, hecho que se refleja en el creciente número de aplicaciones del mundo real que la incluyen como formalización del razonamiento del sentido común. En este sentido, los sistemas argumentativos proponen una formalización de este tipo de razonamiento justamente utilizando como mecanismo de inferencia la argumentación rebatible. Intuitivamente, la argumentación rebatible provee formas de confrontar declaraciones contradictorias para determinar si alguna afirmación puede ser aceptada o rechazada. Para obtener una respuesta, el proceso de razonamiento argumentativo lleva a cabo una serie de etapas. Una etapa muy importante es la comparación de argumentos en conflicto para decidir qué argumento prevalece; ésto requiere introducir un criterio de preferencia entre argumentos que haga frente a esta situación.

Esta tesis aborda el estudio, diseño y formalización de herramientas computacionales concretas para seleccionar y cambiar el criterio de preferencia entre argumentos que es utilizado por el sistema de Programación Lógica Rebatible (DeLP) requerido para decidir derrotas al analizar ataques entre argumentos. Para lograr esto, se proponen varios servicios de razonamiento basados en DeLP que disponen de distintos criterios y permiten llevar a cabo esta tarea de diferentes maneras. Como parte de la contribución, se propone un servicio que utiliza expresiones condicionales para programar cómo seleccionar el criterio que mejor se ajusta a las preferencias del usuario o a una situación en particular. Por otra parte, en la tesis se aborda también la definición de un servicio con mecanismos que permiten no solo seleccionar sino también combinar criterios. Estos mecanismos permiten que sea posible comparar argumentos considerando de manera simultánea más de un criterio.

1. Introducción y Motivación

Los formalismos propuestos en esta tesis incorporan herramientas concretas para tratar el manejo de múltiples criterios de preferencia entre argumentos, lo cual no ha sido considerado hasta el momento por otros trabajos. De esta manera, los resultados obtenidos brindan una contribución importante a los desarrollos en la comunidad de argumentación, particularmente en el campo de los sistemas basados en Programación Lógica Rebatible, significando además un aporte dentro del área de Inteligencia Artificial en las Ciencias de la Computación. En lo que resta de esta sección se presenta brevemente el contexto que motivo el desarrollo de las contribuciones de esta tesis.

1.1. Argumentación Rebatible

Desde hace tiempo, la argumentación ha evolucionado como una propuesta atractiva para modelar el razonamiento basado en sentido común [1, 2, 3], que usualmente sucede en el contexto de información contradictoria, incompleta e incierta. Existen diferentes acercamientos para modelar este proceso, tales como los sistemas argumentativos abstractos [4, 5] que se abstraen de la estructura interna de los argumentos, o los sistemas argumentativos estructurados [6, 7], que sí tienen en cuenta la estructura interna de los mismos. Aquí nos limitamos a estos últimos.

Los sistemas argumentativos estructurados son formalismos de argumentación basados en una lógica subyacente específica, la cual se emplea para representar el conocimiento acerca del dominio sobre el que se razona. Además, esta lógica cuenta con un conjunto de reglas de inferencia que permiten construir argumentos relacionados con una conclusión. Estos sistemas son de gran interés para la comunidad de Inteligencia Artificial dado que las reglas de inferencia permiten representar conocimiento de sentido común, posibilitando la construcción computacional de argumentos. Los sistemas argumentativos estructurados poseen características que los hacen especialmente aptos para su implementación; por otra parte, los sistemas argumentativos son particularmente atractivos para la toma de decisiones y la negociación [8, 9], áreas de gran interés en diversas aplicaciones.

1.2. Programación Lógica Rebatible y la Comparación entre Argumentos

Como se ha mencionado, la importancia del uso de sistemas argumentativos como mecanismo de razonamiento en sistemas inteligentes es reconocida en diversos trabajos en la literatura [2, 10, 11]. Esta tesis se enfoca en un sistema argumentativo estructurado, denominado Programación Lógica Rebatible (DeLP por sus siglas en inglés de *Defeasible Logic Programming*) [7]. Este formalismo combina resultados de programación en lógica y argumentación rebatible, y ha sido aplicado exitosamente en diferentes dominios (ver *e.g.*, [12, 13, 14]). DeLP extiende a la Programación en Lógica permitiendo representar conocimiento potencialmente contradictorio, mediante el uso de la negación fuerte y conocimiento tentativo.

El procedimiento de prueba de DeLP se basa en un proceso de análisis dialéctico donde interactúan argumentos a favor y en contra de un literal a fin de determinar si ese literal está garantizado. Este proceso conducirá a la construcción de una estructura arbórea de derrotadores. De este modo, un literal se hallará garantizado si existe un argumento para ese literal que sobreviva a todas las derrotas que recibe en su árbol asociado. Para determinar si un argumento es un derrotador de otro con el que está en conflicto es necesario de un criterio de preferencia que decida preferencia entre estos argumentos.

Si bien al presentarse DeLP en [7] se asoció el criterio de Especificidad Generalizada como criterio por defecto, en DeLP el criterio es un elemento modular. Sin embargo, en la comunidad de argumentación no existe un consenso establecido acerca de qué criterio utilizar para evaluar argumentos. En este sentido, en trabajos como [15, 16, 17] los autores declaran que el criterio de especificidad no corresponde a un criterio general del razonamiento de sentido común, sino que simplemente es uno de muchos métodos de comparación que podrían ser utilizados. Más aún, propuestas como [18, 15] sugieren que la información acerca del dominio suele ser la herramienta principal para evaluar los argumentos del sistema. De allí la importancia de contar con un sistema como [7] que permita tratar la comparación de argumentos de forma modular, permitiendo que el usuario emplee el criterio que mejor se ajusta al dominio de aplicación reemplazando el provisto por defecto como parte del sistema.

Existen varios criterios concretos para comparar argumentos en DeLP (ver *e.g.*, [7, 19]). No obstante, los sistemas que utilizan DeLP como mecanismo de razonamiento y consideran estos criterios en sus formalismos [20, 21, 22, 23, 24], por lo general adoptan un criterio o una combinación fija de criterios (establecida en la configuración del sistema) de acuerdo al dominio que se está representando. Si bien en DeLP la comparación de argumentos es modular, en la literatura existente no se han propuesto mecanismos programables concretos que le permitan al usuario seleccionar y cambiar dinámicamente el criterio de preferencia dependiendo de sus preferencias o necesidades. Si un usuario conoce las razones por las cuales se prioriza cierta información sobre otra, el mecanismo para comparar argumentos deja de ser una caja negra, aumentando la confianza del usuario sobre las respuestas que puede recibir de estos sistemas en particular. Así el sistema se vuelve más confiable y transparente para el usuario ya que éste puede interactuar directamente con el mecanismo de razonamiento de dicho sistema.

Otra característica de los sistemas basados en DeLP, respecto al manejo de los mecanismos de comparación sobre argumentos, es que usualmente tampoco disponen de varios criterios para que el usuario pueda elegir el que más se adecua a un contexto determinado. En este sentido, en la vida real el proceso de razonamiento humano hace frente a una situación particular teniendo en cuenta diferentes criterios. Por ejemplo, supongamos una persona que se quiere hospedar en un hotel y consulta un sitio web que le ofrece información respecto a varios hoteles. Para elegir un hotel, esta persona podría considerar varios criterios que le permitan comparar la información que le brinda el sitio. Un posible criterio podría preferir aquella información que favorece al confort de las habitaciones, mientras que otro criterio podría preferir la información que favorece a la seguridad del lugar en donde se encuentra el hotel. Finalmente, un tercer criterio podría ser uno que combine los dos anteriores, es decir un criterio que priorice la información relacionada al confort y la seguridad. Por lo tanto, disponer de diferentes criterios de preferencias introduce un grado de flexibilidad adicional a los sistemas basados en DeLP. Sin embargo, y como se puntualizó anteriormente, en estos sistemas el criterio es un componente fijo, o si existe multiplicidad de criterios, no hay forma de cambiarlo una vez que uno es seleccionado e integrado al intérprete. De esta manera, una de las motivaciones de esta tesis es mejorar las capacidades del mecanismo de inferencia de un sistema basado en argumentación para que se pueda adaptar de una forma natural a varios criterios, a través de un mecanismo programable concreto.

A partir de lo mencionado anteriormente, las contribuciones de esta tesis conciernen a la formalización de varios servicios de razonamiento basados en DeLP que abordan los temas planteados en esta sección. En esta dirección, sería posible lograr una satisfactoria integración de razonadores basados en DeLP en diferentes dominios de aplicación, tales como agentes autónomos, sistemas de soporte a las decisiones, búsqueda inteligente en la web, sistemas de recomendación, y otros dominios de similar importancia.

2. Aportes a la disciplina

Como aportes principales de esta tesis se proponen tres servicios de razonamiento basados en DeLP: uno de ellos es el *Servicio de Razonamiento Rebatible basado en Preferencias*, y sus dos extensiones el *Servicio de Razonamiento Rebatible basado en Preferencias Condicionales* y el *Servicio de Razonamiento Rebatible basado en Preferencias Combinadas*. Los resultados de las contribuciones de esta tesis han sido publicados, con referato nacional e internacional, en:

- “*Improving argumentation-based recommender systems through context-adaptable selection criteria*” publicado en la revista *Expert Systems with Applications* vol. 42.
- “*An Approach to Argumentative Reasoning Servers with Multiple Preference Criteria*” publicado en la revista *Inteligencia artificial: Revista Iberoamericana de Inteligencia Artificial* vol. 53.
- “*Servicios de Razonamiento con Múltiples Criterios de Preferencia*” publicado en *XVI Workshop de Investigadores en Ciencias de la Computación (WICC 2014)*.
- “*An Approach to Argumentative Reasoning Servers with Conditions based Preference Criteria*” publicado en *XIX Congreso Argentino de Ciencias de la Computación (CACIC 2013)*.
- “*Modelo de Servicio de Razonamiento con Preferencias*” publicado en el *XV Workshop de Investigadores en Ciencias de la Computación (WICC 2013)*.

3. Contribuciones de la tesis

El objetivo de esta línea de investigación es mejorar las capacidades de razonamiento efectivo de sistemas basados en programación lógica rebatible introduciendo diferentes herramientas programables concretas que permitan tratar el manejo de multiplicidad de criterios. En términos generales, en este trabajo se presentan tres diferentes servicios de razonamiento. A continuación los presentaremos brevemente.

3.1. Servicio de Razonamiento Rebatible basado en Preferencias

En esta sección se presenta la noción de *Servicio de Razonamiento Rebatible basados en Preferencias (SRPref)*, el cual representa el primer aporte de esta tesis. Estos servicios proveen un razonamiento lógico rebatible que cuenta con la capacidad de modificar las preferencias sobre la información procesada cambiando, a partir de cada consulta recibida, el criterio de preferencia utilizado. En particular, el mecanismo que se utiliza para consultar un *Servicio de Razonamiento Rebatible basado en Preferencias (SRPref)* es el de las *Consulta basada en Preferencias*. Entonces, un servicio de razonamiento rebatible basado en preferencias está conformado por:

- el *intérprete* DeLP que se encarga de procesar las consultas hechas al servicio; y
- un *módulo de comparación de argumentos* que, a través de la *función de cómputo de preferencias*, obtiene las preferencias entre argumentos que el intérprete necesita resolver utilizando para ello una de las *implementaciones de criterios de preferencia* almacenadas en el servicio.

Un *SRPref* corresponde a una entidad con la capacidad de responder consultas DeLP utilizando el criterio de preferencia entre argumentos que el usuario especifica junto a su consulta y el programa consultado. Un *SRPref* cuenta con la implementación de varios criterios disponibles para el usuario. Por otra parte, como herramienta de interacción entre el usuario y un *SRPref* se introdujo el concepto de *Consulta basada en Preferencias (CPref)*. Una *CPref* se compone además de la consulta propiamente dicha, del programa DeLP que será procesado y una especificación declarativa del criterio de preferencia que el usuario desea utilizar. En la *especificación de criterio* también se puede incluir información adicional que el criterio elegido podría llegar a necesitar para poder ejecutarse. Por lo tanto, todos estos elementos serán los que un *SRPref* tendrá en cuenta al momento de computar la respuesta a una consulta. Es importante destacar que la *especificación de criterio* constituye la herramienta de interacción que utilizarán todos los servicios de razonamientos propuestos en esta tesis para seleccionar y cambiar de criterio.

En la Figura 1, las flechas muestran cómo son utilizados los datos por el *SRPref*. Entonces, a partir de un programa DeLP y una consulta DeLP, el intérprete calcula la respuesta para la consulta, y el *módulo de comparación de argumentos* computa las preferencias entre argumentos en conflicto que recibe del intérprete. La función de cómputo de preferencias se encargará de ejecutar la implementación del criterio indicado en la especificación de criterio de la consulta. Note que la notación gráfica en Figura 1 para los argumentos es mediante triángulos.

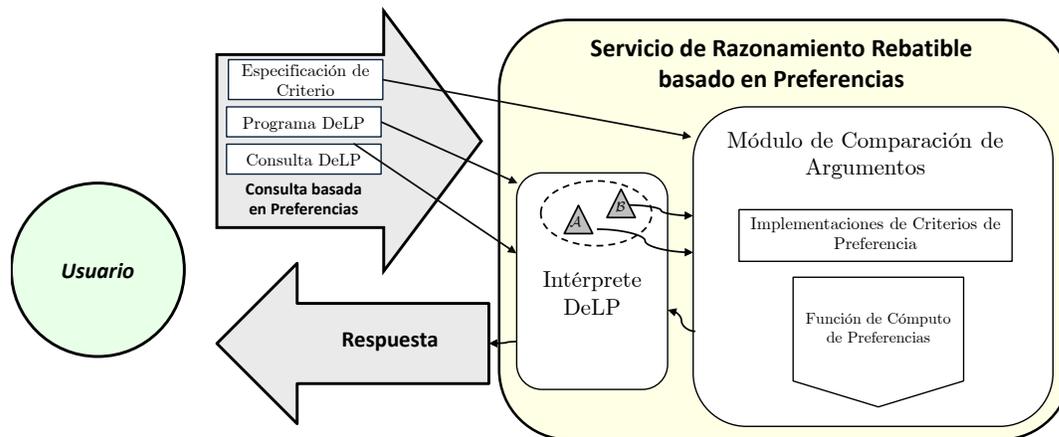


Figura 1: Respuesta a una Consulta basada en Preferencias.

3.2. Servicio de Razonamiento Rebatible basado en Preferencias Condicionales

En primer lugar se presentó el concepto de *Consulta basada en Preferencias Condicionales (CPCCond)* que ofrece una alternativa distinta respecto a la forma en la cuál un usuario puede expresar sus preferencias. Estas consultas están constituidas por una expresión formal que permite decidir cuál es el criterio de preferencia que debe ser utilizado en cada situación específica. Para resolver estas consultas se formalizó un nuevo tipo de servicio de razonamiento que extiende las capacidades de los *SRPrefs*, denominado *Servicio de Razonamiento Rebatible basado en Preferencias Condicionales (SRPCCond)*. Estos servicios tienen la capacidad de responder consultas seleccionando un criterio de preferencia que es indicado utilizando una

expresión condicional incluida en la propia consulta.

Entonces, un *SRPCond* está conformado por:

- el *intérprete* DeLP que resuelve las consultas DeLP,
- una *función de evaluación* que, a partir de una *expresión de preferencia condicional*, obtiene la *especificación de criterio* que declara la implementación de criterio que se utilizará en la comparación de argumentos, y
- un *módulo de comparación de argumentos* que computa las preferencias entre argumentos considerando para ello la especificación de criterio que se obtiene desde la función de evaluación.

Dada la importancia de las *expresiones de preferencias condicionales* como herramienta computacional concreta para guiar la selección de un criterio se presentó una representación de árbol con el fin de proporcionar una forma de analizar varias propiedades de estas expresiones. Estas propiedades son útiles para identificar cuando una expresión puede ser optimizada y de esta manera evitar la computación de literales redundantes, y además caracterizar cuando ciertos caminos en la expresión no son transitables. Estas propiedades son de especial interés en esta tesis por que permiten construir expresiones válidas; es decir, expresiones que mantienen relaciones coherentes entre las guardas que justifican la elección de un criterio en particular.

Un *SRPComb* responde *CPCConds*. Una *CPCCond* se caracteriza por incluir junto a la consulta y el programa a consultar, la especificación de una expresión condicional que servirá para programar cómo seleccionar el criterio que el *SRPCond* consultado utilizará a partir de los criterios especificados. La selección de un criterio en particular depende de la existencia de determinada información alojada en el programa consultado. La Figura 2 muestra de forma esquemática los elementos involucrados en el cálculo de la respuesta para una *CPCCond*. La interpretación de las flechas en el gráfico es la misma que se utilizó en la Figura 1. Por ejemplo, la expresión de preferencia condicional y el programa DeLP son utilizados por la función de evaluación para seleccionar una *especificación de criterio*. Luego, a partir de esta especificación y los argumentos enviados por el intérprete DeLP, el módulo de comparación de argumentos devuelve una preferencia entre dichos argumentos. Finalmente, a partir del programa DeLP y la consulta DeLP, el intérprete obtiene una respuesta para la consulta.

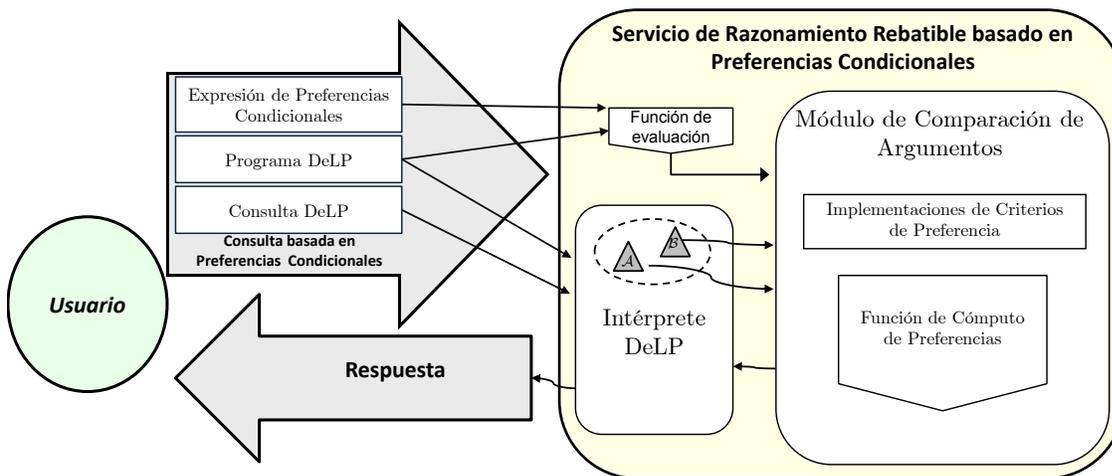


Figura 2: Principales componentes de una *CPCCond* y un *SRPCond*.

3.3. Servicio de Razonamiento Rebatible basado en Preferencias Combinadas

Otro tipo de consultas propuesto en este trabajo son las *Consultas basadas en Preferencias Combinadas (CPComb)*, caracterizadas por incluir una expresión que permite combinar la especificaciones de varios criterios. Para resolver las *CPCombs* se definió un tipo de servicio de razonamiento que extiende las capacidades de un *SRPCond*, denominado *Servicio de Razonamiento Rebatible basado en Preferencias Combinadas (SRPComb)*. Una característica distintiva de estos servicios corresponde a la capacidad de poder usar varios criterios a la vez para evaluar la preferencia entre argumentos. Es decir, estos servicios proveen operadores específicos que permiten a las consultas considerar un uso combinado de varios criterios de preferencia.

Como se muestra de manera esquemática en la Figura 3 los pasos para responder una *CPComb* coinciden en ciertos aspectos con los descriptos para los servicios de razonamiento presentados en secciones anteriores. Es decir, un *SRPComb* cuenta con un intérprete que responde consultas DeLP, y además se comunica con un módulo encargado de resolver las preferencias que éste necesita. Sin embargo, observe que la arquitectura de un *SRPComb* está diseñada de manera tal que puede recibir expresiones que contienen *operadores de combinación de preferencias* utilizados para combinar criterios. Por lo tanto, una vez que la función de evaluación obtiene desde estas expresiones los criterios que se utilizarán, el *módulo extendido de comparación de argumentos* es capaz de llevar a cabo todas las comparaciones entre argumentos que el intérprete le solicita. Para programar la selección de expresiones que incluyan operadores para combinar criterios se formalizó la noción de *expresión de preferencias combinadas*.

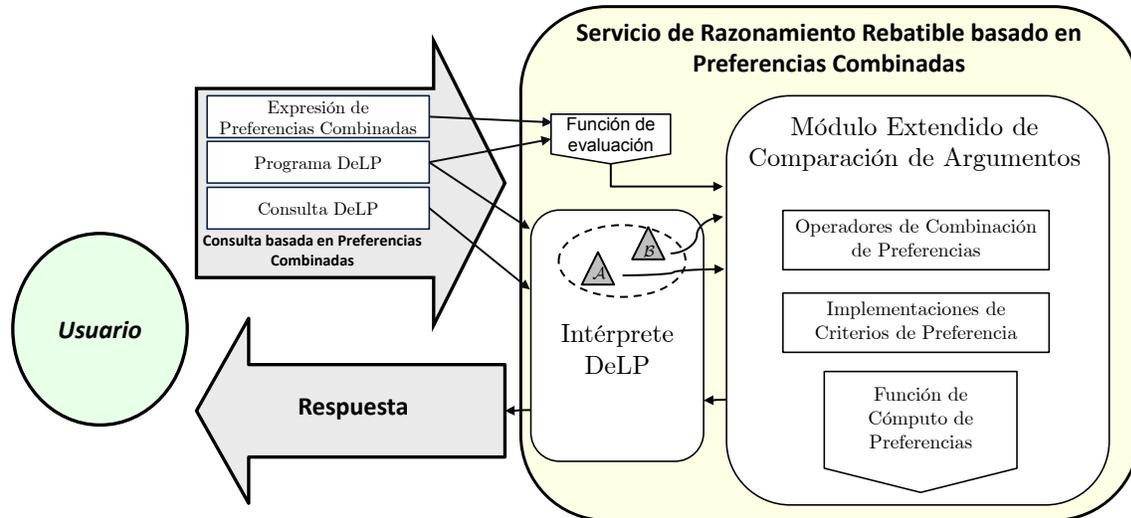


Figura 3: Principales componentes de una *CPComb* y un *SRPComb*.

Se presentaron varios resultados. En primer lugar se estudió un aspecto poco deseable en sistemas argumentativos que es la indecisión por incomparabilidad. Dentro de los resultados mostrados se establecieron algunas propiedades que sirven para establecer cuándo la relación que resulta de aplicar un criterio de preferencia o una operación de combinación de criterios es apropiada a una consulta. Finalmente, se mostraron varios criterios y operadores que mantienen dichas propiedades.

4. Conclusión

A pesar de la importancia que tiene la comparación de argumentos en el proceso de argumentación, los formalismos desarrollados hasta el momento no han considerado mecanismos computacionales que permitan cambiar dinámicamente el criterio de preferencia. Esto motivó el estudio de diferentes herramientas concretas que permitan tratar el manejo de múltiples criterios en aquellos formalismos cuyo mecanismo de inferencia sea DeLP, a fin de poder seleccionar y cambiar de criterio de una manera natural y acorde a las preferencias o necesidades del usuario.

En esta tesis se presentó un servicio de razonamiento basado en la programación lógica rebatible, denominado Servicio de Razonamiento Rebatible basado en Preferencias (*SRPref*), que puede cambiar de criterio de preferencia para cada consulta DeLP recibida. Como herramienta de interacción para consultar estos servicios se introdujo la Consulta basada en Preferencias (*CPref*). Esta herramienta permite especificar de forma declarativa el criterio que será utilizado por el *SRPref* que recibe la consulta.

Se presentó el concepto de Servicio de Razonamiento Rebatible basado en Preferencias Condicionales (*SRPCond*) y se introdujo un tipo especial de consulta para estos servicios, denominada Consulta basada en Preferencias Condicionales (*CPCond*), que le permite a un *SRPCond* seleccionar el criterio que deberá utilizar dependiendo de ciertas condiciones. Para esto la *CPCond* incorpora una expresión de preferencia condicional. Se presentó también una representación de árbol para el estudio de las expresiones condicionales.

Una de las motivaciones importantes en esta tesis ha sido introducir la posibilidad de comparar argumentos considerando varios criterios al mismo tiempo. Siguiendo esta idea se presentó la Consulta basada en Preferencias Combinadas (*CPComb*). Una *CPComb* permite que se pueda especificar el uso de más de un criterio a partir de una expresión especial construida mediante operadores de combinación de preferencias que el servicio consultado dispone.

5. Líneas de investigación futura

Como trabajo futuro se implementarán los desarrollos de esta tesis para poder ejercitar los formalismos y desarrollar nuevas alternativas a partir de las aplicaciones. A partir de los trabajos desarrollados en la tesis se abren varias líneas de investigación sobre las cuales se planea seguir trabajando.

- Se buscará tratar con la formalización de otras herramientas de interacción. Ésto permitirá definir nuevas consultas basadas en preferencias que cumplan con nuevos requisitos del usuario con respecto a sus necesidades o preferencias. Por ejemplo, un usuario podría querer realizar diferentes consultas DeLP simultáneamente utilizando el mismo criterio y el mismo programa DeLP. En este sentido tener un tipo de consulta que cumpla con estos requisitos mejoraría los tiempos de respuesta de los servicios de razonamiento propuestos en esta tesis.
- A partir de las tareas de investigación se observó que la forma en la que es tratada la comparación de argumentos por los formalismos no siempre es la misma. Algunos sistemas utilizan un enfoque modular [7], otros argumentan sobre el criterio de preferencia [6], mientras que otros utilizan un criterio fijo [25]. En este sentido, un tópico interesante es estudiar si es posible construir un marco general para sistemas argumentativos estructurados que permita el manejo de multiplicidad de criterios. Un primer paso para lograr esto,

es analizar la posibilidad de definir servicios de razonamiento en donde el mecanismo de razonamiento sea el de otro sistema argumentativo estructurado, y no el de DeLP. Para esto será necesario diseñar un marco lo suficientemente general como para ser instanciado por diferentes sistemas de este tipo en particular.

- Finalmente, un tema interesante para analizar consiste en el diseño e implementación de una arquitectura de agentes basada en el esquema de alguno de los servicios de razonamiento propuestos.

Referencias

- [1] C. I. Chesñevar, A. G. Maguitman, and R. P. Loui, “Logical models of argument,” *ACM Computing Surveys*, vol. 32, no. 4, pp. 337–383, 2000.
- [2] T. J. M. Bench-Capon and P. E. Dunne, “Argumentation in artificial intelligence,” *Artif. Intell.*, vol. 171, no. 10-15, pp. 619–641, 2007.
- [3] P. Besnard and A. Hunter, *Elements of argumentation*. MIT press Cambridge, 2008, vol. 47.
- [4] P. M. Dung, “On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games,” *Artif. Intell.*, vol. 77, no. 2, pp. 321–358, 1995.
- [5] G. Vreeswijk, “Abstract argumentation systems,” *Artif. Intell.*, vol. 90, no. 1-2, pp. 225–279, 1997.
- [6] H. Prakken and G. Sartor, “Argument-based extended logic programming with defeasible priorities,” *Journal of Applied Non-classical Logics*, vol. 7, pp. 25–75, 1997.
- [7] A. J. García and G. R. Simari, “Defeasible logic programming: An argumentative approach,” *Theory and Practice of Logic Programming (TPLP)*, vol. 4, no. 1-2, pp. 95–138, 2004.
- [8] L. Amgoud, S. Parsons, and N. Maudet, “Arguments, dialogue, and negotiation,” in *ECAI 2000, Proceedings of the 14th European Conference on Artificial Intelligence, Berlin, Germany, August 20-25, 2000*, 2000, pp. 338–342.
- [9] E. Black and A. Hunter, “An inquiry dialogue system,” *Autonomous Agents and Multi-Agent Systems*, vol. 19, no. 2, pp. 173–209, 2009.
- [10] C. I. Chesñevar, A. G. Maguitman, and G. R. Simari, “Recommender system technologies based on argumentation 1,” in *Emerging Artificial Intelligence Applications in Computer Engineering*, 2007, pp. 50–73.
- [11] I. Rahwan and G. R. Simari, *Argumentation in Artificial Intelligence*, 1st ed. Springer Publishing Company, Incorporated, 2009.
- [12] N. D. Rotstein, A. J. García, and G. R. Simari, “Reasoning from desires to intentions: A dialectical framework,” in *AAAI*, 2007, pp. 136–141.

- [13] S. A. Gómez, C. I. Chesñevar, and G. R. Simari, “Defeasible reasoning in web-based forms through argumentation,” *International Journal of Information Technology and Decision Making*, vol. 7, no. 1, pp. 71–101, 2008. [Online]. Available: <http://dx.doi.org/10.1142/S021962200800282X>
- [14] S. Gottifredi, A. J. García, and G. R. Simari, “Query-based argumentation in agent programming,” in *Advances in Artificial Intelligence - IBERAMIA 2010, 12th Ibero-American Conference on AI, Bahía Blanca, Argentina, November 1-5, 2010. Proceedings*, 2010, pp. 284–295.
- [15] G. Vreeswijk, “The feasibility of defeat in defeasible reasoning,” in *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR’91). Cambridge, MA, USA, April 22-25, 1991.*, 1991, pp. 526–534.
- [16] J. L. Pollock, *Cognitive Carpentry: A Blueprint for How to Build a Person*. MIT Press, 1995.
- [17] H. Prakken and G. Sartor, “A dialectical model of assessing conflicting arguments in legal reasoning,” in *Logical Models of Legal Argumentation*. Springer, 1996, pp. 175–211.
- [18] K. Konolige, “Defeasible argumentation in reasoning about events.” in *ISMIS*, 1988, pp. 380–390.
- [19] E. Ferretti, M. Errecalde, A. J. García, and G. R. Simari, “Decision rules and arguments in defeasible decision making,” in *Computational Models of Argument: Proceedings of COMMA 2008, Toulouse, France, May 28-30, 2008.*, 2008, pp. 171–182.
- [20] C. A. D. Deagustini, S. E. F. Dalibón, S. Gottifredi, M. A. Falappa, and G. R. Simari, “Consistent query answering using relational databases through argumentation,” in *Database and Expert Systems Applications - 23rd International Conference, DEXA 2012, Vienna, Austria, September 3-6, 2012. Proceedings, Part II*, 2012, pp. 1–15.
- [21] L. H. Tamargo, S. Gottifredi, A. J. García, M. A. Falappa, and G. R. Simari, “Deliberative delp agents with multiple informants,” *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial*, vol. 15, no. 49, pp. 13–30, 2012.
- [22] L. Godo, E. Marchioni, and P. Pardo, “Extending a temporal defeasible argumentation framework with possibilistic weights,” in *Logics in Artificial Intelligence - 13th European Conference, JELIA 2012, Toulouse, France, September 26-28, 2012. Proceedings*, 2012, pp. 242–254.
- [23] C. E. Briguez, M. Capobianco, and A. G. Maguitman, “A theoretical framework for trust-based news recommender systems and its implementation using defeasible argumentation,” *International Journal on Artificial Intelligence Tools*, vol. 22, no. 4, 2013.
- [24] C. E. Briguez, M. C. Budán, C. A. D. Deagustini, A. G. Maguitman, M. Capobianco, and G. R. Simari, “Argument-based mixed recommenders and their application to movie suggestion,” *Expert Systems with Applications*, vol. 41, no. 14, pp. 6467–6482, 2014.
- [25] T. Wakaki, “Preference-based argumentation capturing prioritized logic programming,” in *Argumentation in Multi-Agent Systems - 7th International Workshop, ArgMAS 2010, Toronto, ON, Canada, May 10, 2010 Revised, Selected and Invited Papers*, 2010, pp. 306–325.